# TRANSACTIONS

OF THE

# CAMBRIDGE

# PHILOSOPHICAL SOCIETY

VOLUME XXII.
(1912—1923)



CAMBRIDGE
AT THE UNIVERSITY PRESS
AND SOLD BY
DEIGHTON, BELL AND CO. LTD. AND BOWES AND BOWES, CAMBRIDGE.
CAMBRIDGE UNIVERSITY PRESS, LONDON.

M.DCCCC.XXIII

# ADVERTISEMENT

*The Society as a body is not to be considered responsible for any facts and opinions advanced in the several Papers, which must rest entirely on the credit of their respective Authors.*

The Society takes this opportunity of expressing its grateful acknowledgments to the Syndics of the University Press for their liberality in contributing towards the expense of printing this Volume of the Transactions.



PRINTED IN GREAT BRITAIN

# CONTENTS OF VOLUME XXII.

# I. *On Compound Denumeration.*

By Major P. A. MacMahon, R.A., Sc.D., LL.D., F.R.S.
Honorary Member Cambridge Philosophical Society.

[*Received* May 1, 1912. *Read* May 6, 1912.]

Art. 1. I propose to examine the subject of compound denumeration, otherwise the partitions of multipartite numbers, by a direct application of the Theory of Distributions which was developed by me in the *Proceedings of the London Mathematical Society**. It will be shewn that the actual denumeration may be made to depend upon the theory of the symmetric functions of a single system of quantities. Such a system is

$$\alpha_1,\ \alpha_2,\ \alpha_3,\ \ldots,$$

and I write in the usual notation

$$(1-\alpha_1 x)(1-\alpha_2 x)(1-\alpha_3 x)\ldots = 1 - a_1 x + a_2 x^2 - a_3 x^3 + \ldots$$
$$= \frac{1}{1 + h_1 x + h_2 x^2 + h_3 x^3 + \ldots},$$

so that the quantities $a$ are the elementary symmetric functions and the quantities $h$ the homogeneous product sums of the quantities $\alpha$ of the system respectively. With these functions are associated the differential operators

$$d_s = \partial_{a_s} + a_1 \partial_{a_{s+1}} + a_2 \partial_{a_{s+2}} + \ldots,$$

$$D_s = \frac{(d_1^s)}{s\,!},$$

where $(d_1^s)$ denotes that the linear operator $d_1$ is raised to the $s$th power in symbolic manner so that it denotes *not* the performance of $d_1$, $s$ times in succession, but rather an operator of the order $s$.

I first consider the partitions of a bipartite number $(pq)$ and note, as observed long ago by me†, that the partitions are separable into groups which depend upon the partitions of the unipartite numbers $(p)$, $(q)$ respectively. Thus the partitions of the bipartite number (22) nine in number are separated into four groups:

| Gr (2, 2), | Gr (2, $1^2$), | Gr ($1^2$, 2), | Gr ($1^2$, $1^2$), |
|---|---|---|---|
| (22) | (21 01) | (12 10) | (11 11) |
| (20 02) | (20 01 01) | (10 10 02) | (11 10 01) |
| | | | (10 10 01 01), |

* *Proc. Lond. Math. Soc.* vol. XIX. 1887, "Symmetric Functions and the Theory of Distributions."

† *American Jour. of Math.* vol. XI. 1888, p. 29, "Memoir on a New Theory of Symmetric Functions."

where it is to be observed that supposing the partible number to be $(pq)$ (here $p=q=2$) the $p$ number is in partition (2) in the two first groups and in partition $(1^2)$ in the last two; while the $q$ number is in partition (2) in the first and third groups and in partition $(1^2)$ in the second and fourth. In fact if the numbers $p$, $q$ have $P$, $Q$ partitions respectively the partitions of $(pq)$ are separable into $PQ$ groups for every partition of $p$ may be associated with every partition of $q$.

We will now study the enumeration of the partitions appertaining to a given group. Consider the group

$$\text{Gr}\,\{(p_1^{\pi_1}p_2^{\pi_2}\ldots),\ (q_1^{\chi_1}q_2^{\chi_2}\ldots)\},$$

where $(p_1^{\pi_1}p_2^{\pi_2}\ldots)$, $(q_1^{\chi_1}q_2^{\chi_2}\ldots)$ are given partitions of $p$ and $q$.

The most extended partition of the group contains $\Sigma\pi+\Sigma\chi$ parts, while that which is least extended contains a number of parts equal to the greatest of the integers $\Sigma\pi$, $\Sigma\chi$.

No generality is lost by the supposition $\Sigma\pi \geqslant \Sigma\chi$.

In a partition of the group the biparts may be ordered so that, as regards the partition of $p$, the first $\Sigma\pi$ biparts are

$$(p_1\,.)\ \pi_1 \text{ times } (p_2\,.)\ \pi_2 \text{ times, \&c.,}$$

and this will be the case for every partition of the group.

The second element of any bipart may be either zero or one of the parts $q_1, q_2, \ldots$ of the partition $(q_1^{\chi_1}q_2^{\chi_2}\ldots)$ of $q$. There may be also biparts of the form $(0q_1)$, $(0q_2)$, .... The biparts are therefore of one of the three forms $(p_sq_t)$, $(p_s0)$, $(0q_t)$, and their number has lower and upper limits $\Sigma\pi$ and $\Sigma\pi+\Sigma\chi$.

We now suppose there to be

$\pi_1$ parcels of one kind
$\pi_2$ " a second kind
&c.
$\Sigma\chi$ parcels of another kind differing from those above.

Altogether $\Sigma\pi+\Sigma\chi$ parcels of a specification which may be denoted by the partition $(\pi_1\pi_2\ldots\Sigma\chi)$ of the number $\Sigma\pi+\Sigma\chi$.

We also suppose there to be

$\chi_1$ objects of one kind
$\chi_2$ " a second kind
&c.
$\Sigma\pi$ objects of another kind differing from those above.

Altogether $\Sigma\pi+\Sigma\chi$ objects of a specification which may be denoted by the partition $(\chi_1\chi_2\ldots\Sigma\pi)$ of the number $\Sigma\pi+\Sigma\chi$.

The number of objects is equal to the number of parcels and we may consider the number of ways of distributing the objects in the parcels so that each parcel contains one object. When, as in the present case, the number of objects is equal to the number of parcels and one object goes into each parcel the notion of the parcel is not essential and we may consider two sets of objects of specifications

$$(\pi_1\pi_2\ldots\Sigma\chi),\quad (\chi_1\chi_2\ldots\Sigma\pi) \text{ respectively;}$$

and the problem is the enumeration of the sets of two-fold objects that can be formed by making $\Sigma\pi + \Sigma\chi$ pairs of objects, each pair consisting of an object from each set of objects.

This problem is precisely the same as that of determining the number of partitions of the bipartite number $(pq)$ which appertain to the group

$$\text{Gr}\{(p_1^{\pi_1} p_2^{\pi_2} \ldots), \quad (q_1^{\chi_1} q_2^{\chi_2} \ldots)\}.$$

To explain this consider the partitions of the bipart (33) which appertain to the group

$$\text{Gr}\{(21),\ (1^3)\}.$$

Here $\pi_1 = 1,\ \pi_2 = 1,\ \Sigma\pi = 2,\ \chi_1 = 3,\ \Sigma\chi = 3.$

We consider objects specified by (113) as being the *first* elements of biparts; these are

$$2,\quad 1,\quad 0,\quad 0,\quad 0;$$

or if we want to exhibit the fact that they are first elements we may write them

$$2*,\quad 1*,\quad 0*,\quad 0*,\quad 0*.$$

With these consider objects specified by (32) as being the second elements of biparts; these are

$$1,\quad 1,\quad 1,\quad 0,\quad 0;$$

or as they may be written

$$*1,\quad *1,\quad *1,\quad *0,\quad *0.$$

Combining the two sets of objects in all possible ways so as to form a single set of two-fold objects we obtain the four sets

$$\begin{array}{ccccc} 21, & 11, & 01, & 00, & 00, \\ 21, & 10, & 01, & 01, & 00, \\ 20, & 11, & 01, & 01, & 00, \\ 20, & 10, & 01, & 01, & 01, \end{array}$$

corresponding to the four partitions

$$\begin{array}{l} (21,\ \ 11,\ \ 01), \\ (21,\ \ 10,\ \ 01,\ \ 01), \\ (20,\ \ 11,\ \ 01,\ \ 01), \\ (20,\ \ 10,\ \ 01,\ \ 01,\ \ 01), \end{array}$$

of the bipartite number (33) appertaining to the group

$$\text{Gr}\{(21),\ (1^3)\}.$$

It is clear that there is in every case a one-to-one correspondence between the distributions as defined and the partitions under examination.

The number of the distributions was shewn (*loc. cit.*) to have either of the two expressions

$$D_{\pi_1} D_{\pi_2} \ldots D_{\Sigma\chi}\, h_{\chi_1} h_{\chi_2} \ldots h_{\Sigma\pi}.$$

$$D_{\chi_1} D_{\chi_2} \ldots D_{\Sigma\pi}\, h_{\pi_1} h_{\pi_2} \ldots h_{\Sigma\chi}.$$

The expressions are equivalent and may be evaluated by means of theorems given (*loc. cit.*).

The whole number of partitions of the bipart $(pq)$ is in consequence

$$\sum_{\pi}\sum_{\chi} D_{\pi_1} D_{\pi_2} \dots D_{\Sigma\chi} h_{\chi_1} h_{\chi_2} \dots h_{\Sigma\pi},$$

the double summation being for every partition $(p_1^{\pi_1} p_2^{\pi_2} \dots)$ of $(p)$ and for every partition $(q_1^{\chi_1} q_2^{\chi_2} \dots)$ of $(q)$.

We apply the method to find the number of partitions of the bipartite (33). We have

| Group | $\pi_1$ | $\pi_2$ | $\Sigma\chi$ | $\chi_1$ | $\chi_2$ | $\Sigma\pi$ | | | Partitions in Group |
|---|---|---|---|---|---|---|---|---|---|
| {(3), (3)} | 1 | 0 | 1 | 1 | 0 | 1 | $D_1^2 h_1^2$ | = | 2 |
| {(3), (21)} | 1 | 0 | 2 | 1 | 1 | 1 | $D_2 D_1 h_1^3$ | = | 3 |
| {(3), ($1^3$)} | 1 | 0 | 3 | 3 | 0 | 1 | $D_3 D_1 h_3 h_1$ | = | 2 |
| {(21), (3)} | 1 | 1 | 1 | 1 | 0 | 2 | $D_1^3 h_2 h_1$ | = | 3 |
| {(21), (21)} | 1 | 1 | 2 | 1 | 1 | 2 | $D_2 D_1^2 h_2 h_1^2$ | = | 7 |
| {(21), ($1^3$)} | 1 | 1 | 3 | 3 | 0 | 2 | $D_3 D_1^2 h_3 h_2$ | = | 4 |
| {($1^3$), (3)} | 3 | 0 | 1 | 1 | 0 | 3 | $D_3 D_1 h_3 h_1$ | = | 2 |
| {($1^3$), (21)} | 3 | 0 | 2 | 1 | 1 | 3 | $D_3 D_2 h_3 h_1^2$ | = | 4 |
| {($1^3$), ($1^3$)} | 3 | 0 | 3 | 3 | 0 | 3 | $D_3^2 h_3^2$ | = | 4 |
| | | | | | | | | Total | 31 |

No calculation is required if we are given Tables which express the $h$ products in terms of monomial symmetric functions.

Thus since a Table shews

$$h_2 h_1^2 = \dots + 7\,(21^2) + \dots$$

$$D_2 D_1^2 h_2 h_1^2 = 7.$$

The above resulting numbers are all shewn in the Tables which proceed as far as the weight 6.

In the above case where $p = q$, it is not necessary to consider all the groups because the two partitions that define the Group may be interchanged. Thus the two Groups {(3), (21)}, {(21), (3)} are identical and the whole numbers of partitions might have been written

$$D_1^2 h_1^2 + 2 D_2 D_1 h_1^3 + 2 D_3 D_1 h_3 h_1 + D_2 D_1^2 h_2 h_1^2 + 2 D_3 D_1^2 h_3 h_2 + D_3^2 h_3^2.$$

We now remark that, formally and algebraically but not operationally, this expression may be written in the factorized form

$$(D_1 h_1 + D_1^2 h_2 + D_3 h_3)(D_1 h_1 + D_2 h_1^2 + D_3 h_3);$$

for the multiplication gives

$$D_1^2 h_1^2 + D_1 D_2 h_1^3 + D_1 D_3 h_1 h_3 + D_1^3 h_2 h_1 + D_1^2 D_2 h_2 h_1^2 + D_1^2 D_3 h_2 h_3 + D_3 D_1 h_3 h_1 + D_3 D_2 h_3 h_1^2 + D_3^2 h_3^2,$$

and, observing that, by the well-known theorem of reciprocity

$$D_2 D_1 h_1^3 = D_1^3 h_2 h_1$$

$$D_3 D_1^2 h_3 h_2 = D_3 D_2 h_3 h_1^2,$$

the truth of the statement is verified.

In fact, formally and algebraically but not operationally, the double sum

$$\sum_\pi \sum_\chi D_{\pi_1} D_{\pi_2} \ldots D_{\Sigma\chi} h_{\chi_1} h_{\chi_2} \ldots h_{\Sigma\pi},$$

may be written

$$\{\sum_\pi D_{\pi_1} D_{\pi_2} \ldots h_{\Sigma\pi}\} \cdot \{\sum_\chi D_{\Sigma\chi} h_{\chi_1} h_{\chi_2} \ldots\}.$$

The factorized form may be regarded as a symbolical expression.

By the above method the following numbers have been calculated

| | 1 | 2 | 3 | 4 | 5 | ... = $q$ |
|---|---|---|---|---|---|---|
| 1 | 2 | — | — | — | — | |
| 2 | 4 | 9 | — | — | — | |
| 3 | 7 | 16 | 31 | — | — | |
| 4 | 12 | 29 | 57 | 109 | — | |
| 5 | 19 | 47 | 97 | 189 | 336 | |
| ⋮ = $p$ | | | | | | |

Thus, from the table, to find the number of partitions of the bipartite (43) we take the row commencing 4 and the column headed 3 and find at the intersection the number 57.

The numbers agree with those obtained by expansion of the generating function

$$\frac{1}{(1-x)(1-y)(1-x^2)(1-xy)(1-y^2)(1-x^3)(1-x^2y)(1-xy^2)(1-y^3)\ldots}.$$

Art. 2. The distribution of $\Sigma\pi + \Sigma\chi$ objects of type $(\pi_1\pi_2 \ldots \Sigma\chi)$ into $\Sigma\pi + \Sigma\chi$ parcels of type $(\chi_1\chi_2 \ldots \Sigma\pi)$ one object in each parcel has necessarily resulted in our obtaining the whole of the partitions of the group under view. Remembering that $\Sigma\pi \geqslant \Sigma\chi$ we may if we please make a distribution of $\Sigma\pi + s$ objects of type $(\pi_1\pi_2 \ldots s)$ into $\Sigma\pi + s$ parcels of type $(\chi_1\chi_2 \ldots \Sigma\pi - \Sigma\chi + s)$ where $s$ is any number included in the series

$$0, \quad 1, \quad 2, \ldots \Sigma\chi.$$

We will thus obtain a number for the enumeration which is

$$D_{\pi_1} D_{\pi_2} \ldots D_s h_{\chi_1} h_{\chi_2} \ldots h_{\Sigma\pi - \Sigma\chi + s};$$

and this number also gives the number of partitions of $(pq)$ which contain $\Sigma\pi + s$ or fewer parts and also appertain to the given group.

Hence also the number of such partitions which contain exactly $\Sigma\pi + s$ parts is

$$D_{\pi_1} D_{\pi_2} \ldots D_s h_{\chi_1} h_{\chi_2} \ldots h_{\Sigma\pi - \Sigma\chi + s} - D_{\pi_1} D_{\pi_2} \ldots D_{s-1} h_{\chi_1} h_{\chi_2} \ldots h_{\Sigma\pi - \Sigma\chi + s - 1};$$

or as it may be written

$$D_{\pi_1} D_{\pi_2} \ldots \{D_s h_{\chi_1} h_{\chi_2} \ldots h_{\Sigma\pi - \Sigma\chi + s} - D_{s-1} h_{\chi_1} h_{\chi_2} \ldots h_{\Sigma\pi - \Sigma\chi + s - 1}\}.$$

The whole number of partitions which contain $\Sigma\pi + s$ or fewer parts is

$$\sum_\pi \sum_\chi D_{\pi_1} D_{\pi_2} \ldots D_s h_{\chi_1} h_{\chi_2} \ldots h_{\Sigma\pi - \Sigma\chi + s},$$

the double summation being for all partitions

$$(p_1^{\pi_1} p_2^{\pi_2} \ldots),\ (q_1^{\chi_1} q_2^{\chi_2} \ldots) \text{ of the numbers } (p) \text{ and } (q);$$

and a similar summation gives the whole number of partitions which contain exactly $\Sigma\pi + s$ parts. The expression for the number of partitions which contain $\Sigma\pi + s$ or fewer parts can be given the factorized symbolic form

$$(\sum_\pi D_{\pi_1} D_{\pi_2} \ldots h_{\Sigma\pi - \Sigma\chi + s}) \cdot (\sum_\chi D_s h_{\chi_1} h_{\chi_2} \ldots).$$

As an example let us consider the partitions of the bipartite (44) which appertain to the group {(211), (211)}.

Here $s$ may have the values 0, 1, 2, 3.

$$\pi_1 = 1,\ \pi_2 = 2,\ \Sigma\chi = 3,\ \chi_1 = 1,\ \chi_2 = 2,\ \Sigma\pi = 3.$$

For $s = 0$, we have $D_2 D_1 h_2 h_1 = 2$, and the two partitions into 3 parts are

(22, 11, 11),
(21, 12, 11).

For $s = 1$, we have $D_2 D_1^2 h_2 h_1^2 = 7$, shewing that there are 7 partitions into 4 or fewer parts; in addition to the 2 which have exactly 3 parts already written down we have 5 which contain exactly 4 parts; these are

(22, 11, 10, 01),
(21, 12, 10, 01),
(21, 11, 10, 02),
(20, 12, 11, 01),
(20, 11, 11, 02).

For $s = 2$, we have $D_2^2 D_1 h_2^2 h_1 = 11$, and we find that in addition to the 7 forms already written we have 4 which contain exactly 5 parts; these are

(22 10 10 01 01),
(21 10 10 02 01),
(20 12 10 01 01),
(20 11 10 02 01).

Finally for $s = 3$ we have $D_3 D_2 D_1 h_3 h_2 h_1 = 12$, and we have $12 - 11 = 1$ partition which contains exactly 6 parts; this is

(20 10 10 02 01 01).

* To explain the general method of calculation it is to be noted that

$$D_s h_m = h_{m-s},$$

and that when operating upon a product, $D_s$ acts through each of the partitions of $s$. Thus

$$D_2 h_l h_m = h_{l-2} h_m + h_l h_{m-2} + h_{l-1} h_{m-1}$$

$$\begin{aligned} D_3 h_l h_m h_n = {} & h_{l-3} h_m h_n + h_l h_{m-3} h_n + h_l h_m h_{n-3} \\ & + h_{l-2} h_{m-1} h_n + h_{l-2} h_m h_{n-1} + h_l h_{m-2} h_{n-1} \\ & + h_{l-1} h_{m-2} h_n + h_{l-1} h_m h_{n-2} + h_l h_{m-1} h_{n-2} \\ & + h_{l-1} h_{m-1} h_{n-1}. \end{aligned}$$

* *Vide Proc. Lond. Math. Soc.* vol. xix. 1887, pp. 127—128, "The Algebra of Multi-linear partial differential operators."

The calculation of $D_3D_2D_1 . h_3h_2h_1$ therefore proceeds as follows:—

$$D_3D_2D_1 . h_3h_2h_1 = D_3D_2(h_2^2h_1 + h_3h_1^2 + h_3h_2)$$
$$= D_3(2h_2h_1 + h_1^3 + 2h_2h_1 + h_1^3 + h_3 + 2h_2h_1 + h_2h_1 + h_3 + h_2h_1)$$
$$= D_3(2h_1^3 + 8h_2h_1 + 2h_3) = 2 + 8 + 2 = 12.$$

Art. 3. In the next place we examine the effect of employing the elementary functions $a_1, a_2, a_3, \ldots$ instead of the homogeneous product sums $h_1, h_2, h_3, \ldots$. The distributions enumerated by the number

$$D_{\pi_1}D_{\pi_2}\ldots D_s a_{\chi_1}a_{\chi_2}\ldots a_{\Sigma\pi-\Sigma\chi+s},$$

are those of objects of type $(\pi_1\pi_2\ldots s)$ into parcels of type $(\chi_1\chi_2\ldots\Sigma\pi-\Sigma\chi+s)$ one object being placed in each parcel subject to the restriction that no two similar objects are to be placed in similar parcels.

The corresponding partitions of the bipartite number $(pq)$ are those which appertain to the group $\{(p_1^{\pi_1}p_2^{\pi_2}\ldots), (q_1^{\chi_1}q_2^{\chi_2}\ldots)\}$, which contain exactly $\Sigma\pi+s$ parts, the zero bipart 00 not being excluded as a permissible bipart, and in which no particular part (including the bipart 00) occurs more than once.

Of course the double sum

$$\sum_{\pi}\sum_{\chi} D_{\pi_1}D_{\pi_2}\ldots D_s a_{\chi_1}a_{\chi_2}\ldots a_{\Sigma\pi-\Sigma\chi+s},$$

enumerates such partitions for the totality of the groups.

To see the meaning of this result consider again the partitions of the bipartite number (44) which appertain to the group $\{(211), (211)\}$.

For $s=0$, we have $D_2D_1a_2a_1=1$; since $\Sigma\pi=3$, this means that of all the partitions of the group which contain exactly 3 parts, the zero part 00 being admissible, there is but one in which there are no similarities of parts. This partition is in fact

(21 12 11).

For $s=1$, we have $D_2D_1^2a_2a_1^2=5$; for a Table which expresses $a$ products in terms of monomial symmetric functions gives

$$a_2a_1^2 = \ldots + 5(21^2) + \ldots,$$

giving $$D_2D_1^2a_2a_1^2 = 5.$$

Thence we conclude that there are just 5 partitions which contain 4 parts involving no similarities. These are

(21 12 11 00),
(22 11 10 01),
(21 12 10 01),
(21 11 10 02),
(20 12 11 01),

the set including the one previously found with the part 00 added.

For $s=2$, we have $D_2^2 D_1 a_2^2 a_1 = 5$, since

$$a_2^2 a_1 = \ldots + 5\,(2^2 1) + \ldots;$$

and the 5 forms indicated are found to be

$$\begin{array}{lllll} (22 & 11 & 10 & 01 & 00), \\ (21 & 12 & 10 & 01 & 00), \\ (21 & 11 & 10 & 02 & 00), \\ (20 & 12 & 11 & 01 & 00), \\ (20 & 11 & 10 & 02 & 01), \end{array}$$

the parts in each involving no similarities.

Finally for $s=3$, we have $D_3 D_2 D_1 a_3 a_2 a_1 = 1$, since

$$a_3 a_2 a_1 = \ldots + (321) + \ldots$$

and the form indicated is

$$(20 \quad 11 \quad 10 \quad 02 \quad 01 \quad 00)$$

containing no similarities of parts.

We obtain information concerning the partitions of the group which contain different parts when 00 is excluded as a part; for denote by $Q_s$ the numbers of partitions of the group which contain exactly $s$ different parts, the zero part being excluded, we have

$$Q_3 = 1,\ Q_3 + Q_4 = 5,\ Q_4 + Q_5 = 5,\ Q_5 = 1,$$

whence $Q_4 = 4$ and $Q_3 + Q_4 + Q_5 = 6$.

This number 6 which enumerates the partitions of the group which possess different parts is either

$$D_3 D_2 D_1 a_3 a_2 a_1 + D_2 D_1^2 a_2 a_1^2,$$

or

$$D_2^2 D_1 a_2^2 a_1 + D_2 D_1 a_2 a_1.$$

In general we have the relations

$$Q_{\Sigma\pi} = D_{\pi_1} D_{\pi_2} \ldots a_{\chi_1} a_{\chi_2} \ldots a_{\Sigma\pi - \Sigma\chi},$$

$$Q_{\Sigma\pi} + Q_{\Sigma\pi+1} = D_{\pi_1} D_{\pi_2} \ldots D_1 a_{\chi_1} a_{\chi_2} \ldots a_{\Sigma\pi - \Sigma\chi + 1},$$

$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots$$

$$Q_{\Sigma\pi + \Sigma\chi - 2} + Q_{\Sigma\pi + \Sigma\chi - 1} = D_{\pi_1} D_{\pi_2} \ldots D_{\Sigma\chi - 1} a_{\chi_1} a_{\chi_2} \ldots a_{\Sigma\pi - 1},$$

$$Q_{\Sigma\pi + \Sigma\chi - 1} = D_{\pi_1} D_{\pi_2} \ldots D_{\Sigma\chi} a_{\chi_1} a_{\chi_2} \ldots a_{\Sigma\pi}.$$

Hence the number which enumerates those partitions of the group which have different parts, the zero part being excluded, has two expressions; for

$$\begin{aligned} Q_{\Sigma\pi} &+ Q_{\Sigma\pi+1} + \ldots + Q_{\Sigma\pi + \Sigma\chi - 1} \\ &= D_{\pi_1} D_{\pi_2} \ldots a_{\chi_1} a_{\chi_2} \ldots a_{\Sigma\pi - \Sigma\chi} + D_{\pi_1} D_{\pi_2} \ldots D_2 a_{\chi_1} a_{\chi_2} \ldots a_{\Sigma\pi - \Sigma\chi + 2} \\ &\qquad + D_{\pi_1} D_{\pi_2} \ldots D_4 a_{\chi_1} a_{\chi_2} \ldots a_{\Sigma\pi - \Sigma\chi + 4} + \ldots \\ &= D_{\pi_1} D_{\pi_2} \ldots D_1 a_{\chi_1} a_{\chi_2} \ldots a_{\Sigma\pi - \Sigma\chi + 1} + D_{\pi_1} D_{\pi_2} \ldots D_3 a_{\chi_1} a_{\chi_2} \ldots a_{\Sigma\pi - \Sigma\chi + 3} \\ &\qquad + D_{\pi_1} D_{\pi_2} \ldots D_5 a_{\chi_1} a_{\chi_2} \ldots a_{\Sigma\pi - \Sigma\chi + 5} + \ldots, \end{aligned}$$

where if $\Sigma\chi$ be uneven both series extend to $\frac{1}{2}(\Sigma\chi + 1)$ terms, whilst if $\Sigma\chi$ be even the first and second series extend to $\frac{1}{2}\Sigma\chi$ and $\frac{1}{2}\Sigma\chi - 1$ terms respectively.

That these two series are equivalent may be shewn algebraically as follows.

For brevity put $\Sigma\pi - \Sigma\chi = \theta$ and note that

$$D_1 a_{\chi_1} a_{\chi_2} \dots a_{\theta+1} = a_{\chi_1} a_{\chi_2} \dots a_\theta + \Sigma a_{\chi_1 - 1} a_{\chi_2} \dots a_{\theta+1},$$

$$D_2 a_{\chi_1} a_{\chi_2} \dots a_{\theta+2} = \Sigma a_{\chi_1 - 1} a_{\chi_2} \dots a_{\theta+1} + \Sigma a_{\chi_1 - 1} a_{\chi_2 - 1} a_{\chi_3} \dots a_{\theta+2},$$

$$D_3 a_{\chi_1} a_{\chi_2} \dots a_{\theta+3} = \Sigma a_{\chi_1 - 1} a_{\chi_2 - 1} a_{\chi_3} \dots a_{\theta+2} + \Sigma a_{\chi_1 - 1} a_{\chi_2 - 1} a_{\chi_3 - 1} a_{\chi_4} \dots a_{\theta+3},$$

.......................................................................................

&c.

Directly we operate upon the relations with $D_{\pi_1} D_{\pi_2} \dots$ the equivalence is obvious.

Art. 4. There is no difficulty in extending this theory by filling up the gap between the elementary functions and the homogeneous product sums. For suppose $k_1, k_2, k_3, \dots$ be functions derived from the homogeneous product sums by deleting therefrom all terms which involve quantities of the system (from which the symmetric functions are derived) to a higher power than $k$. Then the distributions enumerated by the number

$$D_{\pi_1} D_{\pi_2} \dots D_s k_{\chi_1} k_{\chi_2} \dots k_{\Sigma\pi - \Sigma\chi + s},$$

are those of objects of type $(\pi_1 \pi_2 \dots s)$ into parcels of type $(\chi_1 \chi_2 \dots \Sigma\pi - \Sigma\chi + s)$ one object being placed in each parcel subject to the restriction that more than $k$ similar objects are not to be placed in similar parcels. The corresponding partitions of the bipartite number $(pq)$ are those which appertain to the group $\{(p_1^{\pi_1} p_2^{\pi_2} \dots), (q_1^{\chi_1} q_2^{\chi_2} \dots)\}$, which contain exactly $\Sigma\pi + s$ parts, the zero part 00 not being excluded from being an admissible part, and in which no particular part (including the part 00) occurs more than $k$ times.

Art. 5. I pass on to consider the similar theory of tripartite partitions and it will be found to shew what the theory is for multipartite partitions in general. Consider the tripartite number $(pqr)$ and the partitions appertaining to the group

$$\{(p_1^{\pi_1} p_2^{\pi_2} \dots), (q_1^{\chi_1} q_2^{\chi_2} \dots), (r_1^{\rho_1} r_2^{\rho_2} \dots)\},$$

wherein we will suppose $\Sigma\pi \geqslant \Sigma\chi \geqslant \Sigma\rho$.

The partitions involve at least $\Sigma\pi$ and at most $\Sigma\pi + \Sigma\chi + \Sigma\rho$ parts. Reasoning as in the bipartite case we find that for partitions into $\Sigma\pi + s$ parts, where $0 \leqslant s \leqslant \Sigma\chi + \Sigma\rho$, we have to do with three assemblages of objects of types

$$(\pi_1 \pi_2 \dots s),\ (\chi_1 \chi_2 \dots t),\ (\rho_1 \rho_2 \dots u) \text{ respectively,}$$

where

$$\Sigma\pi + s = \Sigma\chi + t = \Sigma\rho + u.$$

We have to consider the number of ways of forming $\Sigma\pi + s$ triads of objects by taking one object from each assemblage to form a triad.

This number is also the number of partitions, appertaining to the group, which involve $\Sigma\pi + s$ or fewer parts.

This problem in 'Distributions' was solved by me in *American Journal of Mathematics,* vol. XIV. 1892, pp. 33 *et seq.,* "Fourth Memoir on a New Theory of Symmetric Functions."

Let there be two systems of quantities

$$\alpha_1, \quad \alpha_2, \quad \alpha_3, \ \ldots$$
$$\beta_1, \quad \beta_2, \quad \beta_3, \ \ldots$$

and let their symmetric functions be denoted by partitions with suffixes 1 and 2 respectively. Then write

$$A_1 = (1)_1$$
$$A_2 = (2)_1 + (1^2)_1$$
$$A_3 = (3)_1 + (21)_1 + (1^3)_1$$
$$\cdots\cdots\cdots\cdots\cdots\cdots$$
$$B_1 = (1)_2 A_1$$
$$B_2 = (2)_2 A_2 + (1^2)_2 A_1^2$$
$$B_3 = (3)_2 A_3 + (21)_2 A_2 A_1 + (1^3)_2 A_1^3$$
$$\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots$$

where it will be noted that $A_1, A_2, A_3, \ldots$ are the successive homogeneous product sums of the quantities $\alpha_1, \alpha_2, \alpha_3, \ldots$*.

We now form the product

$$B_{\pi_1} B_{\pi_2} \ldots B_s,$$

and eliminate the quantities $A_1, A_2, A_3, \ldots$ so as to express it as a linear function of terms each of which is a product of two symmetric functions denoted by partitions with suffixes 1 and 2 respectively.

One of these terms will be

$$M(\chi_1\chi_2 \ldots t)_1 (\rho_1\rho_2 \ldots u)_2,$$

where $M$ is an integer which is equal to the number of distributions in question.

We have therefore to find the coefficient of

$$(\chi_1\chi_2 \ldots t)_1 (\rho_1\rho_2 \ldots u)_2,$$

in the development of the product

$$B_{\pi_1} B_{\pi_2} \ldots B_s.$$

Let $D_1^{(1)}, D_2^{(1)}, D_3^{(1)} \ldots; D_1^{(2)}, D_2^{(2)}, D_3^{(2)}, \ldots$, be obliterating operators associated with symmetric functions of the quantities $\alpha_1, \alpha_2, \alpha_3, \ldots; \beta_1, \beta_2, \beta_3, \ldots$ respectively. Then the operators $D^{(1)}, D^{(2)}$, act upon functions which are denoted by partitions in brackets $(\ )_1, (\ )_2$, respectively.

From the well-known properties of these operators we know that

$$D_{\chi_1}^{(1)} D_{\chi_2}^{(1)} \ldots D_t^{(1)} . D_{\rho_1}^{(2)} D_{\rho_2}^{(2)} \ldots D_u^{(2)} . B_{\pi_1} B_{\pi_2} \ldots B_s = M.$$

The reader will have no difficulty in establishing that

$$D_s^{(1)} A_m = A_{m-s},$$
$$D_s^{(2)} B_m = A_s B_{m-s},$$

* These quantities were denoted by $h_1, h_2, h_3, \ldots$ in Art. 2.

relations which much facilitate the calculation of the number $M$. To take a simple example, consider the partitions of the tripartite number (333) which appertain to the group

$$\{(21), (21), (21)\}.$$

Here
$$\pi_1 = 1,\ \pi_2 = 1,\ \chi_1 = 1,\ \chi_2 = 1,\ \rho_1 = 1,\ \rho_2 = 1,$$
$$t = u = s,$$

and $s$ may have the values 0, 1, 2, 3, 4, for

$$s = 0, \quad (D_1^{(1)})^2 (D_1^{(2)})^2 B_1^2 = 4,$$
$$s = 1, \quad (D_1^{(1)})^3 (D_1^{(2)})^3 B_1^3 = 36,$$
$$s = 2, \quad (D_1^{(1)})^2 (D_2^{(1)}) (D_1^{(2)})^2 (D_2^{(2)}) B_1^2 B_2 = 74,$$
$$s = 3, \quad (D_1^{(1)})^2 (D_3^{(1)}) (D_1^{(2)})^2 (D_3^{(2)}) B_1^2 B_3 = 86,$$
$$s = 4, \quad (D_1^{(1)})^2 (D_4^{(1)}) (D_1^{(2)})^2, (D_4^{(2)}) B_1^2 B_4 = 87\ ;$$

shewing that the number of partitions of the group which contain

exactly 2 parts is 4,
„ 3 „ 32,
„ 4 „ 38,
„ 5 „ 12,
„ 6 „ 1;

and of course the total number of the partitions of the group is 87.

To explain the above calculations the reader is reminded that $D_s^{(1)}$ and $D_s^{(2)}$ operate through the whole of the partitions of $s$ upon a $B$ product. Thus for example

$$\begin{aligned}
D_3^{(2)} B_s B_t B_u = & (D_3^{(2)} B_s) B_t B_u + B_s (D_3^{(2)} B_t) B_u + B_s B_t (D_3^{(2)} B_u) \\
& + (D_2^{(2)} B_s)(D_1^{(2)} B_t) B_u + (D_2^{(2)} B_s) B_t (D_1^{(2)} B_u) + B_s (D_2^{(2)} B_t)(D_1^{(2)} B_u) \\
& + (D_1^{(2)} B_s)(D_2^{(2)} B_t) B_u + (D_1^{(2)} B_s) B_t (D_2^{(2)} B_u) + B_s (D_1^{(2)} B_t)(D_2^{(2)} B_u) \\
& + (D_1^{(2)} B_s)(D_1^{(2)} B_t)(D_1^{(2)} B_u) \\
= & A_3 (B_{s-3} B_t B_u + B_s B_{t-3} B_u + B_s B_t B_{u-3}) \\
& + A_2 A_1 (B_{s-2} B_{t-1} B_u + B_{s-2} B_t B_{u-1} + B_s B_{t-2} B_{u-1} \\
& + B_{s-1} B_{t-2} B_u + B_{s-1} B_t B_{u-2} + B_s B_{t-1} B_{u-2}) \\
& + A_1^3 B_{s-1} B_{t-1} B_{u-1},
\end{aligned}$$

where the partitions of 3 being (3), (21), ($1^3$) the first line, the next two lines, and the fourth line are given by the three partitions respectively. The result

$$(D_1^{(1)})^2 (D_4^{(1)}) . (D_1^{(2)})^2 (D_4^{(2)}) . B_1^2 B_4 = 87,$$

is obtained as follows :—

$$\begin{aligned}
(D_1^{(1)})^2 (D_4^{(2)}) . B_1^2 B_4 &= (D_1^{(2)}) (D_4^{(2)}) (2A_1 B_1 B_4 + A_1 B_1^2 B_3) \\
&= A_1 D_4^{(2)} (2A_1 B_4 + 2A_1 B_1 B_3 + 2A_1 B_1 B_3 + A_1 B_1^2 B_2) \\
&= A_1^2 (2A_4 + 4A_1 A_3 + A_1^2 A_2)\ ;
\end{aligned}$$

therefore $(D_1^{(1)})^2 (D_4^{(1)}) . (D_1^{(2)})^2 (D_4^{(2)}) . B_1^2 B_4$

$$\begin{aligned}
&= (D_1^{(1)}) (D_4^{(1)}) (4A_1 A_4 + 2A_1^2 A_3 + 12A_1^2 A_3 + 4A_1^3 A_2 + 4A_1^3 A_2 + A_1^5) \\
&= (D_1^{(1)}) (D_4^{(1)}) (4A_1 A_4 + 14A_1^2 A_3 + 8A_1^3 A_2 + A_1^5) \\
&= (D_4^{(1)}) (4A_4 + 4A_1 A_3 + 28A_1 A_3 + 14\, A_1^2 A_2 + 24A_1^2 A_2 + 8A_1^4 + 5A_1^4) \\
&= (D_4^{(1)}) (4A_4 + 32A_1 A_3 + 38A_1^2 A_2 + 13A_1^4) \\
&= 4 + 32 + 38 + 13 = 87.
\end{aligned}$$

We have found that the number of partitions of the group which have $\Sigma\pi + s$ or fewer parts is equal to the coefficient of

$$(\chi_1\chi_2 \dots t)_1 (\rho_1\rho_2 \dots u)_2,$$

in the development of the product

$$B_{\pi_1} B_{\pi_2} \dots B_s.$$

By a well known theorem of symmetry we may in this theorem interchange in any manner the partitions

$$(\pi_1\pi_2 \dots s),\ (\chi_1\chi_2 \dots t),\ (\rho_1\rho_2 \dots u).$$

We may therefore carry out the calculation in 3! different ways; a circumstance that is convenient for the purpose of verification.

The total number of partitions of the tripartite $(pqr)$ is

$$\sum_\chi \sum_\rho \sum_\pi D_{\chi_1}^{(1)} D_{\chi_2}^{(1)} \dots D_t^{(1)} . D_{\rho_1}^{(2)} D_{\rho_2}^{(2)} \dots D_u^{(2)} . B_{\pi_1} B_{\pi_2} \dots B_s,$$

the summation being for the whole of the partitions of $p$, $q$ and $r$.

Art. 6. We have also the theory of the partitions of the group which are composed of different parts; it is merely necessary in the above to substitute for the homogeneous product sums $A_1, A_2, A_3, \dots$ the elementary functions $a_1, a_2, a_3, \dots$.

Thus in the above particular case

$$(D_1^{(1)})^2 . (D_1^{(2)})^2 . B_1^2$$
$$= (D_1^{(1)})^2 . (2a_1^2) = D_1^{(1)} . 4a_1 = 4,$$
$$(D_1^{(1)})^3 . (D_1^{(2)})^3 . B_1^3$$
$$= (D_1^{(1)})^3 . (6a_1^3) = (D_1^{(1)})^2 . 18a_1^2 = 36,$$
$$(D_1^{(1)})^2 (D_2^{(1)}) . (D_1^{(2)})^2 (D_2^{(2)}) . B_1^2 B_2$$
$$= (D_1^{(1)})^2 (D_2^{(1)}) . (5a_1^4 + 2a_1^2 a_2) = (D_1^{(1)})^2 (34a_1^2 + 2a_2) = 70,$$
$$(D_1^{(1)})^2 (D_3^{(1)}) . (D_1^{(2)})^2 (D_3^{(2)}) . B_1^2 B_3$$
$$= (D_1^{(1)})^2 (D_3^{(1)}) . (a_1^5 + 4a_1^3 a_2 + 2a_1^2 a_3) = (D_1^{(1)})^2 (22a_1^2 + 6a_2) = 50,$$
$$(D_1^{(1)})^2 (D_4^{(1)}) . (D_1^{(2)})^2 (D_4^{(2)}) . B_1^2 B_4$$
$$= (D_1^{(1)})^2 (D_4^{(1)}) . (a_1^4 a_2 + 4a_1^3 a_3 + 2a_1^2 a_4) = (D_1^{(1)})^2 (4a_1^2 + 5a_2) = 13.$$

For a given number of parts we take all the corresponding partitions of the group zero parts 000 being admissible as parts; then the numbers found indicate the number of partitions which have no repeated parts.

Thus of 2 parts there are 4 partitions in which no part is repeated

| | | | | |
|---|---|---|---|---|
| 3 | ,, | 36 | ,, | ,, |
| 4 | ,, | 70 | ,, | ,, |
| 5 | ,, | 50 | ,, | ,, |
| 6 | ,, | 13 | ,, | ,, |

Moreover if $Q_s$ denote the number of partitions of the groups which contain exactly $s$ different parts, zero parts excluded,

$$Q_2 = 4,$$
$$Q_2 + Q_3 = 36,$$
$$Q_3 + Q_4 = 70,$$
$$Q_4 + Q_5 = 50,$$
$$Q_5 + Q_6 = 13;$$

whence $$Q_2 = 4,\ Q_3 = 32,\ Q_4 = 38,\ Q_5 = 12,\ Q_6 = 1.$$

Comparing these numbers with those found in Art. 6 we see that there are in fact no partitions, of the group, which involve repeated parts.

Art. 7. The theory in respect of multipartite numbers in general is now clear. We take the multipartite number $(pqrs\ldots)$ and the partitions appertaining to the group

$$\{(p_1^{\pi_1} p_2^{\pi_2}\ldots),\ (q_1^{\chi_1} q_2^{\chi_2}\ldots),\ (r_1^{\rho_1} r_2^{\rho_2}\ldots),\ (s_1^{\sigma_1} s_2^{\sigma_2}\ldots)\ldots\}.$$

We continue the series of relations

$$A_1 = (1)_1 \qquad B_1 = (1)_2 A_1 \qquad C_1 = (1)_3 B_1 \qquad \&\text{c.}\ldots \qquad M_1 = (1)_{n-1} L_1$$
$$A_2 = (2)_1 + (1^2)_1 \quad B_2 = (2)_2 A_2 + (1^2)_2 A_1^2, \quad C_2 = (2)_3 B_2 + (1^2)_3 B_1^2 \qquad M_2 = (2)_{n-1} L_2 + (1^2)_{n-1} L_1^2$$
$$\&\text{c.} \qquad \&\text{c.} \qquad \&\text{c.}$$

where if the multipartite number be $n$-partite, $L$, $M$ are the $n-2$th and $n-1$th letters of the alphabet. We have then to find the coefficient of

$$(\chi_1\chi_2\ldots t_2)_1 (\rho_1\rho_2\ldots t_3)_2 (\sigma_1\sigma_2\ldots t_4)_3 \ldots$$

in the development of $M_{\pi_1} M_{\pi_2} \ldots M_{t_1}$. We have the sought number equal to

$$D_{\chi_1}^{(1)} D_{\chi_2}^{(1)} \ldots D_{t_2}^{(1)} . D_{\rho_1}^{(2)} D_{\rho_2}^{(2)} \ldots D_{t_3}^{(2)} . D_{\sigma_1}^{(3)} D_{\sigma_2}^{(3)} \ldots D_{t_4}^{(3)} \ldots\ldots M_{\pi_1} M_{\pi_2} \ldots M_{t_1};$$

and observe that

$$D_s^{(3)} . C_m = B_s C_{m-s},$$
$$D_s^{(4)} . D_m = C_s D_{m-s},$$
$$\vdots$$
$$D_s^{(n-1)} M_m = L_s M_{m-s},$$

relations which enable the regular and progressive calculations of the sought number.

In all cases the theory of the partitions into dissimilar parts is reached by substituting the elementary functions $a_1, a_2, a_3, \ldots$ for the homogeneous product sums $A_1, A_2, A_3, \ldots$.

The totality of the partitions into $\Sigma\pi + t_1$ or fewer parts is given by the expression

$$\sum_\chi \sum_\rho \sum_\sigma \ldots D_{\chi_1}^{(1)} D_{\chi_2}^{(1)} \ldots D_{t_2}^{(1)} . D_{\rho_1}^{(2)} D_{\rho_2}^{(2)} \ldots D_{t_3}^{(2)} . D_{\sigma_1}^{(3)} D_{\sigma_2}^{(3)} \ldots D_{t_3}^{(3)} \ldots\ldots M_{\pi_1} M_{\pi_2} M_{\pi_3} \ldots M_{t_1},$$

the summation being in regard to all the partitions

$$(p_1^{\pi_1} p_2^{\pi_2}\ldots),\ (q_1^{\chi_1} q_2^{\chi_2}\ldots),\ (r_1^{\rho_1} r_2^{2\rho}\ldots),\ (s_1^{\sigma_1} s_2^{\sigma_2}\ldots),\ \ldots$$

of the numbers $p, q, r, s, \ldots$ respectively.

## II. *A class of integral functions defined by Taylor's series.*

By G. N. WATSON, M.A.

[*Received* Sept. 20, 1912. *Read* Nov. 11, 1912.]

1. IT is sometimes possible to determine the complete asymptotic expansion of an integral function by making use of asymptotic expansions of more simple integral functions. An instance is afforded by the deduction* of the complete asymptotic expansion of the function

$$F_\beta(x;\theta)=\sum_{n=0}^{\infty}\frac{x^n\,.\,\chi(n+\theta)}{n!\,(n+\theta)^\beta}$$

[where $\chi(y)$ is analytic† in the vicinity of $y=\infty$] from the asymptotic expansions of functions of the type

$$G_\beta(x;\theta)=\sum_{n=0}^{\infty}\frac{x^n}{n!\,(n+\theta)^\beta}.$$

In this memoir, I propose to obtain the asymptotic expansion of a class of integral functions of a more general nature than the function $F_\beta(x;\theta)$ defined above. I am inclined to think that the integral functions which will be considered are the most general integral functions which possess the two properties (i) that the coefficient of the $n$th term in the Taylor's series which defines the function is a simple function of $n$, and (ii) that the asymptotic expansions of the functions involve only powers and exponentials of the variable.

2. Let $f(x)$ be a function defined by Taylor's series

$$f(x)=c_0+c_1x+c_2x^2+\ldots, \quad \text{.................................(1)},$$

and let it be possible to define a function $\phi(s)$, which is analytic in certain regions (to be specified presently) of the plane of the complex variable $s$, such that when $s$ is equal to any positive integer $n$,

$$\frac{e^{gn}\phi(n)}{\Gamma(an+1)}=c_n,$$

where $\phi(n)\,.\exp\{\beta'\log n\}\rightarrow$ a finite limit as $n\rightarrow\infty$; in order that $f(x)$ may be an integral function we must take

$$R(a)>0;$$

* Barnes, *Phil. Trans. Roy. Soc.*, vol. CCVI. A. (1906), pp. 273—278.

† There is a large class of functions $F_\beta(x;\theta)$ which are such that $\chi(y)$ is not analytic for large values of $y$, although $\chi(y)$ possesses an asymptotic expansion of the form

$$\chi(y)=b_0+b_1y^{-1}+b_2y^{-2}+\ldots,$$

when $y$ is large and real; to such functions Barnes' results do not apply.

apart from this restriction, $\alpha$, $\beta'$, $g$ are any constants, real or complex; it is convenient to determine $\alpha$ so that $|\arg \alpha| < \frac{1}{2}\pi$.

The hypotheses* which we shall now make concerning the function $\phi(s)$ are as follows:

(i) That a number, $h_1$, exists such that $\phi(s)$ is analytic on the right of the line

$$R(\alpha s) = h_1.$$

(ii) That numbers $\gamma'$ and $\lambda$ exist such that $\gamma' > 0$, $0 < \lambda < \frac{1}{2}\pi$ and such that when $|s| \geqslant \gamma'$, $|\arg(s/\alpha)| \leqslant \lambda + \frac{1}{2}\pi$, then $\phi(s/\alpha) . \exp(\beta' \log s)$ possesses an asymptotic expansion† in negative powers of $s$, the grade‡ and outer grade‡ of the expansion being equal to 1; that is to say, that when $|s| \geqslant \gamma'$, $|\arg(s/\alpha)| \leqslant \lambda + \frac{1}{2}\pi$, then $\phi(s/\alpha)$ can be expanded into a series

$$\phi(s/\alpha) = \exp(-\beta' \log s) . \left[a_0'' + \frac{a_1''}{s} + \frac{a_2''}{s^2} + \dots + \frac{a_n''}{s^n} + R_n\right],$$

where $|a_n''| < A_1 . \rho^n . n!$, $|R_n s^{n+1}| < A_2 . \sigma_1^n . n!$, it being supposed that $A_1$, $A_2$, $\rho$ and $\sigma_1$ are independent of $n$.

When $\phi(s)$ is subject to these conditions, I have shewn§ that $\phi(s/\alpha)$ can be expanded into either of the two following series of inverse factorials:

$$\text{(i)} \quad \phi(s/\alpha) = \frac{1}{(s+\Theta)^\beta}\left[a_0 + \frac{a_1}{Ms+1} + \frac{a_2}{(Ms+1)(Ms+2)} + \frac{a_3}{(Ms+1)(Ms+2)(Ms+3)} + \dots\right] \quad \text{......(2a)},$$

$$\text{(ii)} \quad \phi(s/\alpha) = \frac{1}{(s+\Theta)^\beta}\left[a_0' + \frac{a_1'}{\{M(s-\mu)+1\}} + \frac{a_2'}{\{M(s-\mu)+1\}\{M(s-\mu)+2\}} + \frac{a_3'}{\{M(s-\mu)+1\}\{M(s-\mu)+2\}\{M(s-\mu)+3\}} + \dots\right] \quad \text{......(2b)}.$$

In these expansions, $\Theta$ is any complex‖ number such that $-R(\Theta) > \gamma'$; the expansions are valid and the series on the right converge when $R(s+\Theta) > 0$; $\mu = -R(\Theta)$ and $\beta$ is a number¶ such that $R(\beta) < 0$; and if $B$ be the integer such that $0 > R(\beta + B) \geqslant -1$, then the coefficients in the expansions satisfy the inequalities

$$|a_k| < H . \Gamma(k) . k^{\mu} \{\log(k+1)\}^B \text{..............................(2c)},$$

$$|a_k'| < H . \Gamma(k) . \{\log(k+1)\}^B \text{..................................(2d)},$$

when $k \geqslant 1$ and $H$ is some number independent of $k$; $M$ is any number less than $M_0$ where $M_0$ is a positive number depending on $\rho$ and $\lambda$.

In this memoir, I propose to investigate the asymptotic expansion of $f(x)$, defined by the series (1), for large values of $x$ when the conditions stated above are satisfied by $\phi(s/\alpha)$, so that

* These conditions are satisfied if $\phi$ be a member of a very large class of functions which can easily be constructed. Some examples are given below in § 15.

† A comparison of equation (8) below with Part v. of Barnes' memoir (*loc. cit.*) shews that, in the special case when the expansion for $\phi(s/\alpha)$ is not asymptotic but convergent, the asymptotic expansion of the function $f(x)$ can easily be obtained by Barnes' methods.

‡ These terms are introduced in a memoir by the writer "A theory of asymptotic series," *Phil. Trans. Roy. Soc.*, vol. CCXI. A. (1911), pp. 279—313.

§ See *Rendiconti del Circolo Matematico di Palermo*, t. XXXIV. pp. 65—84.

‖ $\Theta$ must not be purely real.

¶ If $R(\beta') < 0$ we take $\beta = \beta'$; if $R(\beta') \geqslant 0$ we choose $\beta$ so that $\beta' - \beta$ is a positive integer.

the expansions (2$a$) and (2$b$) are valid when $R(s+\Theta)>0$; these two expansions will be utilised in obtaining the asymptotic expansion in question.

When $|\alpha-1|>1$, it will be shewn that the asymptotic expansion of $f(x)$ can be obtained for all values of $\arg x$.

When $|\alpha-1|\leqslant 1$, the asymptotic expansion of $f(x)$ can be obtained for a certain range of values of $\arg x$; in the part of the plane not included in this range the asymptotic expansion of $f(x)$ depends on the behaviour of $f(s)$ on the left of the line $R(\alpha s)=h_1$.

The analysis to be employed is so much simpler when we may take $M=1$ than when it is necessary to take $M<1$, that we investigate the case $M=1$ separately in Parts I and II of the paper; the case when $M<1$ is investigated in Parts III and IV.

The symbols* $K$, $O$ and $o$ will be employed throughout to mean 'a definite constant,' 'of the order of,' and 'of order less than,' respectively; thus $f(x)=O(g(x))$ means that $\underset{x\to\infty}{\text{Lt}}\sup\{|f(x)|\div g(x)\}$ is finite; while $f(x)=o(g(x))$ means that $\underset{x\to\infty}{\text{Lt}}\sup\{|f(x)|\div g(x)\}=0$.

In Parts I and III of the paper, $K$ will be supposed to be independent of the variables $x$ and $y$ and also of a variable integer $k$.

## PART I. *Preliminary asymptotic formulae.*

3. Let us define the integral function † $E_k(x)$ by the series,

$$E_k(x)=\sum_{n=0}^{\infty}\frac{x^n}{\Gamma(n+k+1)}\quad\ldots\ldots\ldots(3),$$

where $k$ is a (large) integer; we shall obtain asymptotic expressions for $E_k(x)$ for all values of $x$.

(i) Let $|x|\leqslant 1$; then

$$|E_k(x)|=\frac{1}{\Gamma(k+1)}\left|\left\{1+\frac{x}{k+1}+\frac{x^2}{(k+1)(k+2)}+\ldots\right\}\right|$$

$$<\frac{1}{\Gamma(k+1)}\left\{1+\frac{1}{k+1}+\frac{1}{(k+1)^2}+\ldots\right\}$$

$$<\frac{2}{\Gamma(k+1)}\text{ when } k\geqslant 1.$$

Thus, when $|x|\leqslant 1$ and $k\geqslant 1$,

$$|E_k(x)|<\frac{2}{\Gamma(k+1)}\quad\ldots\ldots\ldots(3\,a).$$

(ii) Let $|x|\geqslant 1$, $|\arg x|\leqslant\frac{1}{2}\pi-\delta$ where $\delta>0$; then by multiplication of series

$$e^{-x}E_k(x)=\frac{1}{\Gamma(k+1)}\sum_{n=0}^{\infty}\frac{(-)^n kx^n}{n!\,(k+n)}=-\frac{1}{2\pi i}\int\frac{1}{\Gamma(k+1)}\frac{k}{s+k}\Gamma(-s)\,x^s\,ds,$$

along a contour which may be taken parallel to the imaginary axis‡ passing through the point $s=-l$; where $l$ is some fixed integer, and $k$ is taken to be such that $k\geqslant l+1$.

On the contour $|k(s+k)^{-1}|\leqslant l+1$ and $\int|\Gamma(-s)\,x^s\,ds|<2\pi K$ where $K$ depends on $l$ and $\delta$ only when $|x|>1$.

Thus, when $|x|>1$, $|\arg x|\leqslant\frac{1}{2}\pi-\delta$, $k\geqslant l+1$

$$|E_k(x)|<\frac{K}{\Gamma(k+1)}\left|\frac{e^x}{x^l}\right|\quad\ldots\ldots\ldots(3\,b).$$

* The use of the symbols $K$ and $O$ is explained by Hardy, "Orders of Infinity" (*Camb. Math. Tracts*, No. 12); the symbol $O$ was introduced by Landau (*Primzahlen*, Bd I. p. 61).

† Some properties of this function have been given by Hardy (*Proc. Lond. Math. Soc.* ser. 2, vol. II. pp. 404—405).

‡ See Barnes' memoir (*loc. cit.*), Part I.

To discuss the function $E_k(x)$ for other values of $x$, we need the following Lemma*:

*Lemma*. If $1 \geqslant u_1 \geqslant u_2 \geqslant \ldots \geqslant 0$, and if $s_n$ denote the sum of $n+1$ terms of the series $1 + u_1 e^{i\omega} + u_2 e^{2i\omega} + \ldots$ $\omega$ being a real angle, then

$$|s_n| \leqslant |\operatorname{cosec}(\tfrac{1}{2}\omega)|.$$

[For $$(1-e^{i\omega}) S_n = 1 + (u_1 - 1) e^{i\omega} + (u_2 - u_1) e^{2i\omega} + \ldots + (u_n - u_{n-1}) e^{ni\omega} - u_n e^{(n+1) i\omega},$$

so that $$|(1-e^{i\omega}) S_n| \leqslant 1 + (1 - u_1) + (u_1 - u_2) + \ldots + (u_{n-1} - u_n) + u_n$$
$$\leqslant 2,$$

i.e. $|S_n| < |\operatorname{cosec}(\frac{1}{2}\omega)|$.]

(iii) Let $x = re^{i\omega}$ where $r \leqslant k$, $\delta \leqslant \omega \leqslant 2\pi - \delta$ and $0 < \delta < \frac{1}{2}\pi$.

Then $$E_k(x) = \frac{1}{\Gamma(k+1)} \left[1 + \frac{r}{k+1} e^{i\omega} + \frac{r^2}{(k+1)(k+2)} e^{2i\omega} + \ldots\right].$$

Applying the lemma, we get

$$|E_k(x)| < \frac{\operatorname{cosec} \frac{1}{2}\delta}{\Gamma(k+1)} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (3c).$$

(iv) Let $x = re^{i\omega}$ where $r \geqslant k$, $\delta \leqslant \omega \leqslant 2\pi - \delta$ and $0 < \delta < \frac{1}{2}\pi$.

Then $$E_k(x) - \frac{e^x}{x^k} = -\frac{1}{x\Gamma(k)} \left[1 + \frac{k-1}{r} e^{-i\omega} + \frac{(k-1)(k-2)}{r^2} e^{-2i\omega} + \ldots\right].$$

Applying the lemma, we get

$$\left|E_k(x) - \frac{e^x}{x^k}\right| < \frac{\operatorname{cosec} \frac{1}{2}\delta}{x\,\Gamma(k)} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (3d).$$

The results $(3a) \ldots (3d)$ are those which will be required concerning $E_k(x)$.

4. Let us now consider the behaviour of the integral function

$$G_\beta(x; \theta, k) = \sum_{n=0}^{\infty} \frac{x^n}{\Gamma(n+k+1) \cdot (n+\theta)^\beta},$$

where $|x|$ is large, $k$ is a (large) integer, $0 < R(\theta) \leqslant 1$ and $R(\beta) < 0$; for convenience we take $|x| \geqslant 2$.

As in the portion of Barnes' memoir† which deals with the function $G_\beta(x, \theta)$ we may shew that

$$G_\beta(x; \theta, k) = \frac{e^{-\pi i\beta}\,\Gamma(1-\beta)}{2\pi i x} \int \left[-\log\left(1 - \frac{y}{x}\right)\right]^{\beta-1} \left(1 - \frac{y}{x}\right)^{\theta-1} E_k(x-y) \,.\, dy,$$

the integral being taken along a contour starting from the point $x$, encircling the origin in a positive direction and returning to the point $x$. This contour is marked with double arrows in Fig. 1. The many-valued functions are specified by taking

$$\arg(1 - y/x) = 0, \quad \arg\{-\log(1 - y/x)\} = 0,$$

when $y$ lies on $xO$ before the circuit of the origin has been made.

Now let us deform the contour into that marked by single arrows in Fig. 1; the four parallel lines make a non-zero angle with the line $Ox$, and they make an angle less than $\frac{1}{3}\pi$ with the real axis; the circle surrounding $O$ is the circle $|y| = 1$; the lines $PQ$ are

* Cp. Bromwich, *Theory of Infinite Series*, § 20.

† We have

$$\frac{1}{(n+\theta)^\beta} = \frac{e^{-\pi i\beta}\,\Gamma(1-\beta)}{2\pi i} \int z^{\beta-1} e^{-(\theta+n)z} dz$$

round a contour starting from $+\infty$, encircling the origin and returning to $+\infty$; we may justify the integration term by term, of the series

$$\sum_{n=0}^{\infty} \frac{z^{\beta-1} e^{-(\theta+n)z} x^n}{\Gamma(n+k+1)},$$

round this contour by Bromwich, *Theory of Infinite Series*, § 176 A, and then put $1 - y/x = e^{-z}$; the function $G_\beta(x, \theta)$ is dealt with in Part III. of the memoir already quoted.

parallel to the imaginary axis and at a distance from it equal to $|x|\{1+k\exp|x|\}$; so that the lengths $xP$ and $OQ$ lie between $k\exp|x|$ and $|x|\{2+k\exp|x|\}\sec\frac{1}{3}\pi$; i.e. if $k\geqslant 1$, $xP$ and $OQ$ are less than $6k|x|\exp|x|$. We note that on the loop from $Q$ round the origin, when $|y|\geqslant 2|x|$, $R(x-y)\leqslant 0$, since $OQ$ is inclined at an angle less than $\frac{1}{3}\pi$ to the real axis.

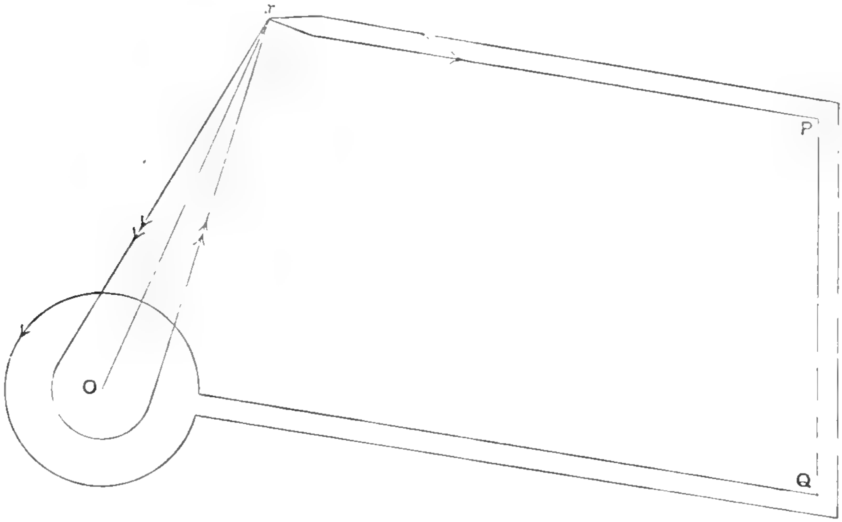



Fig. 1.

We may now write the formula for $G_\beta(x;\theta,k)$ in the form

$$G_\beta(x;\theta,k)=\frac{e^{-\pi i\beta}\Gamma(1-\beta)}{2\pi i x}\int_C\left[\log\left(1-\frac{y}{x}\right)\right]^{\beta-1}\left(1-\frac{y}{x}\right)^{\theta-1}E_k(x-y)\,dy$$
$$-\frac{1}{x\Gamma(\beta)}\int_{PQ}\left[-\log\left(1-\frac{y}{x}\right)\right]^{\beta-1}\left(1-\frac{y}{x}\right)^{\theta-1}E_k(x-y)\,dy$$
$$-\frac{1}{x\Gamma(\beta)}\int_{OP'}\left[-\log\left(-\frac{\eta}{x}\right)\right]^{\beta-1}\left(-\frac{\eta}{x}\right)^{\theta-1}E_k(-\eta)\,d\eta \qquad (4),$$

where $C$ denotes the contour starting from $Q$, encircling the origin, and returning to $Q$; in the second and third integrals the many-valued functions are specified by the values which they had at $P$ in the original loop integral before the circuit of the origin was made; since we have written $y=\eta+x$, $OP'$ is a line equal and parallel to $xP$.

We write this formula for $G_\beta(x;\theta,k)$ in the form

$$G_\beta(x;\theta,k)=I_1+I_2+I_3 \qquad (4a),$$

and we proceed to find inequalities satisfied by $|I_1|$, $|I_2|$, $|I_3|$.

Let us first consider $I_2$; on $PQ$ we have the following inequalities satisfied:

$$\left|\log\left(1-\frac{y}{x}\right)\right|>K,\quad \left|\left(1-\frac{y}{x}\right)\right|>K,\quad \left|\arg\left\{-\log\left(1-\frac{y}{x}\right)\right\}\right|<K,\quad \left|\arg\left(1-\frac{y}{x}\right)\right|<K;$$

also $R(x-y)<0$, so that by (3a), (3c) and (3d) we have $|E_k(x-y)|<K\div\Gamma(k+1)$; and it is easy to see that the length $PQ$ is less than $2|x|$.

Consequently $$|I_2| < \frac{1}{|x\,\Gamma(\beta)|}\int_{PQ} K\{\Gamma(k+1)\}^{-1}\,d\,|y|.$$

Hence $$|I_2| < \frac{K}{\Gamma(k+1)} \quad\text{.................................}(5\,a).$$

We next consider the integral $I_3$: it is easy to shew, by $(3\,c)$ and $(3\,d)$, that on the path of integration

$$E_k(-\eta) < \frac{K}{\Gamma(k)\,.\,\{k+|\eta|\}}.$$

Putting $\eta = -rxe^{i\omega}$ on the path of integration ($r$ being real and positive), we see that

$$|I_3| < \frac{1}{|x\,.\,\Gamma(\beta)|}\int_0^\rho |(-\log r - i\omega)^{\beta-1}|\,.\,|(re^{i\omega})^{\theta-1}|\,.\,\frac{K|x|}{\Gamma(k)\,.\,\{k+r|x|\}}\,dr,$$

where $\rho$ is some number less than $6k\exp|x|$.

Since $|(-\log r - i\omega)| > \frac{1}{2}\{|\log r| + |\omega|\}$, it is easy to see that

$$|I_3| < \frac{K}{\Gamma(k)}\int_0^\rho \frac{r^{R(\theta)-1}\,dr}{\{|\log r| + |\omega|\}^{1-R(\beta)}\{k+r|x|\}}.$$

We now divide the path of integration into three parts, viz.

(i) from $r=0$ to $r=k_1$, where $k_1 = k^{\frac{1}{2}}$,

(ii) from $r=k_1$ to $r=k$,

(iii) from $r=k$ to $r=\rho$.

There are two cases to be considered, the first when $0 < R(\theta) < 1$, the second when $R(\theta) = 1$; in both cases we observe that:

on (i), $\{|\log r| + |\omega|\}^{R(\beta)-1} < K$,

on (ii), $\{|\log r| + |\omega|\}^{R(\beta)-1} < K\{\log k_1\}^{R(\beta)-1}$,

on (iii), $\{|\log r| + |\omega|\}^{R(\beta)-1} < K\{\log k\}^{R(\beta)-1}$.

In the first case, we see that

$$|I_3| < \frac{K}{\Gamma(k)}\left[\int_{(i)}\frac{r^{R(\theta)-1}}{k+r|x|}\,dr + \{\log k_1\}^{R(\beta)-1}\int_{(ii)}\frac{r^{R(\theta)-1}}{k+r|x|}\,dr + \{\log k\}^{R(\beta)-1}\int_{(iii)}\frac{r^{R(\theta)-1}}{k+r|x|}\,dr\right]$$

$$< \frac{K}{\Gamma(k)}\left[\int_{(i)} k^{-1}r^{R(\theta)-1}\,dr + \{\tfrac{1}{2}\log k\}^{R(\beta)-1}\int_{(ii)} k^{-1}r^{R(\theta)-1}\,dr + \{\log k\}^{R(\beta)-1}\int_k^\infty |x|^{-1}r^{R(\theta)-2}\,dr\right]$$

$$< \frac{K}{\Gamma(k)}\left[k^{\frac{1}{2}R(\theta)-1} + k^{R(\theta)-1}\{\log k\}^{R(\beta)-1} + |x|^{-1}k^{R(\theta)-1}\{\log k\}^{R(\beta)-1}\right].$$

So that $$|I_3| < \frac{K}{\Gamma(k+1)}\,\frac{k^{R(\theta)}}{\{\log k\}^{1-R(\beta)}} \quad\text{.................................}(5\,b).$$

[NOTE. In the investigation of $I_3$ we ought, strictly speaking, to have taken the contour in the immediate vicinity of $x$ to be of the form shewn in Fig. 2, since the point $x$ is an essential singularity of the integrand: but it is easy to see that, on the corrected path (i), the inequality $\{|\log r| + |\omega|\}^{R(\beta)-1} < K$ is still true, and thence that the contribution of the corrected path (i) to the integral $I_3$ is still $O(\{\Gamma(k)\}^{-1}k^{\frac{1}{2}R(\theta)-1})$; it being assumed that the arcs of the small circles near $x$ are of radius less than, say, $\frac{1}{2}$ so that they are not in the immediate vicinity of $O$ or $P$.]

In the second case, when $R(\theta)=1$, we have to treat the integral along the path (iii) in a slightly different manner; we have

$$\int_{(iii)} \frac{r^{R(\theta)-1}}{k+r|x|}\,dr = \int_1^{\rho} \frac{dr}{k+r|x|} = \frac{1}{|x|}\log\frac{k+\rho|x|}{k(1+|x|)} < \frac{1}{|x|}\log(\rho/k) < \tfrac{1}{2}(2+\log 6), \text{ since } |x|\geqslant 2:$$

and hence $(5b)$ is still true even when $R(\theta)=1$.

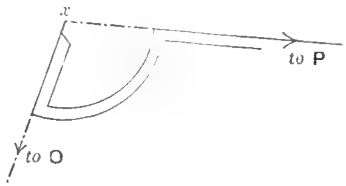


Fig. 2.

We next have to estimate the value of the integral

$$\int_C \left[-\log\left(1-\frac{y}{x}\right)\right]^{\beta-1}\left(1-\frac{y}{x}\right)^{\theta-1} E_k(x-y)\,.\,dy.$$

We divide the path of integration into two portions, the first portion being such that on it $|y|\leqslant 2|x|$, the second being such that on it $|y|\geqslant 2|x|$; we call the integrals along these portions $I_1'$ and $I_1''$ respectively.

On the portion for which $|y|\leqslant 2|x|$, we notice that

$$\left|\frac{x}{y}\log\left(1-\frac{y}{x}\right)\right| > K, \quad \left|\arg\frac{x}{y}+\arg\left\{-\log\left(1-\frac{y}{x}\right)\right\}\right| < K,$$

so that

$$|I_1'| < K\int\left|\left(\frac{y}{x}\right)^{\beta-1}\left(1-\frac{y}{x}\right)^{\theta-1} E_k(x-y)\,.\,dy\right|;$$

while we remember that $|y|\geqslant 1$, and $\left|\left(1-\frac{y}{x}\right)^{\theta-1}\right| < K$, since $R(\theta)\leqslant 1$.

It follows that $$|I_1'| < \frac{K}{|x^{\beta-1}|}\int|E_k(x-y)\,.\,dy|.$$

Now all the asymptotic expressions for $E_k(x-y)$ are comprised in the single formula:

$$|E_k(x-y)| < K\left|\frac{e^{x-y}}{(x-y)^l\,\Gamma(k+1)}\right| + \frac{K}{\Gamma(k+1)};$$

and on the portion of the contour under consideration $|x-y|>K|x|$.

Hence $$|I_1'| < \frac{1}{|x^{\beta-1}|\,.\,\Gamma(k+1)}\int K\{|x^{-l}e^{x-y}|+1\}\,.\,|dy|.$$

Also, on the contour, $|e^{-y}|<K$, while the length of the contour is less than $K|x|$.

Therefore, finally,

$$|I_1'| < \frac{1}{\Gamma(k+1)}\left\{\frac{K\,|e^{x}|}{|x^{l+\beta-2}|}+\frac{K}{|x^{\beta-2}|}\right\} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots(5c)$$

Lastly we have to estimate the value of

$$I_1'' = \int\left[-\log\left(1-\frac{y}{x}\right)\right]^{\beta-1}\left(1-\frac{y}{x}\right)^{\theta-1} E_k(x-y)\,dy,$$

the integration being along those portions of $C$ for which $|y| \geqslant 2|x|$; on these portions of $C$, $R(x-y) \leqslant 0$, so that by (3c) and (3d)

$$|E_k(x-y)| < \frac{K}{\Gamma(k).\{k+|(x-y)|\}}.$$

Putting $1-y/x = re^{i\omega}$ ($r$ real) we divide the path of integration into three portions*, viz. (i) from the point where $y=2x$ to the point where $|1-y/x| = k^{\frac{1}{2}}$; (ii) from this point to the point where $|1-y/x| = k$; (iii) from this last point to the point where $y = OQ$; noticing that $|dy| < K|x|.dr$, we may prove†, in precisely the same way as we proved (5b), that

$$|I_1''| < \frac{K|x|.k^{R(\theta)}}{\Gamma(k+1).(\log k)^{1-R(\beta)}} \quad\text{.............................(5d).}$$

[From Fig. 1 we observe that each of the three portions above has to be taken twice, but that fact does not alter the form of this result.]

Collecting the results numbered (5) we see that, if $l$ be a fixed integer, $k \geqslant l+1$, $|x| > 2$, and $0 < R(\theta) \leqslant 1$, then

$$|G_\beta(x;\theta,k)| < \frac{K|e^x|}{|x^{l+\beta-1}|.\Gamma(k+1)} + \frac{K|x^{1-\beta}|}{\Gamma(k+1)} + \frac{K.k^{R(\theta)}}{\Gamma(k+1).\{\log k\}^{1-R(\beta)}} \quad\text{............(6),}$$

where $K$ is independent of $k$ and $x$.

We wish to extend this result, so as to cover a greater range of values of $\theta$.

First, let $\theta = N + \phi$, where $N$ is a positive integer and $0 < R(\phi) \leqslant 1$.

Then
$$G_\beta(x;\theta,k) = \frac{1}{x^N} G_\beta(x;\phi,k-N) - \sum_{n=0}^{N-1} \frac{x^{n-N}}{\Gamma(n+k+1-N).(n-N+\theta)^\beta}.$$

We easily deduce from (6) that, when $N < R(\theta) \leqslant N+1$ and $k \geqslant N+l+1$,

$$|G_\beta(x;\theta,k)| < \frac{K|e^x.k^N|}{|x^{N+l+\beta-1}|.\Gamma(k+1)} + \frac{K|x^{1-\beta}.k^N|}{\Gamma(k+1)} + \frac{K.k^{R(\theta)}}{\Gamma(k+1).\{\log k\}^{1-R(\beta)}} \quad\text{......(6a).}$$

Similarly, when $-N < R(\theta) \leqslant -N+1$, $N$ being a positive integer and $k \geqslant -N+l+1$,

$$|G_\beta(x;\theta,k)| < \frac{K|x|^N}{\Gamma(k+1)} + \frac{K|e^x|}{|x^{l+\beta-N-1}|\Gamma(N+k+1)} + \frac{K|x^{1-N-\beta}|}{\Gamma(k+N+1)} + \frac{K|x|^N k^{R(\theta)}}{\Gamma(k+1).(\log k)^{1-R(\beta)}} \quad\text{......(6b).}$$

It is convenient to quote the complete asymptotic expansion of $G_\beta(x;\theta,k)$ when $|x|$ is large and $k$ is fixed; we have

$$G_\beta(x;\theta,k) = \frac{1}{x^k} G_\beta(x;\theta-k) - \sum_{n=0}^{k-1} \frac{x^{n-k}}{\Gamma(n+1).(n-k+\theta)^\beta};$$

consequently‡, if $\quad [\log(1-y)]^{\beta-1}(1-y)^{\theta-k-1} \equiv (-y)^{\beta-1} \sum_{n=0}^{\infty} d_n(-y)^n,$

$$G_\beta(x;\theta,k) = -\sum_{n=0}^{k-1} \frac{x^{n-k}}{\Gamma(n+1).(n-k+\theta)^\beta} + \frac{e^x}{x^{k+\beta}}\left[\sum_{n=0}^{N} \frac{(-)^n d_n \Gamma(\beta+n)}{\Gamma(\beta)x^n} + o(x^{-N})\right]$$

$$+ (-x)^{-\theta}[\log(-x)]^{\beta-1}\left[\sum_{n=0}^{N} \frac{(-)^{n+k}\Gamma^{(n)}(\theta-k)}{n!\,\Gamma(\beta-n)[\log(-x)]^n} + o([\log(-x)]^{-N})\right] \quad\text{...(6c),}$$

where $|\arg x| \leqslant \pi$, $|\arg(-x)| \leqslant \pi$ and $N$ is any assigned integer.

* Some of these portions will be incomplete or missing when $k$ is not large.

† The work is not worth setting out in full.

‡ Barnes' memoir, *loc. cit.* Part III.

These form the set of asymptotic formulae which will be sufficient for our purposes at present.

PART II. *The asymptotic expansion of $f(x)$ when $M=1$.*

5. We are now in a position to attack the asymptotic expansion of $f(x)$ as defined in § 2. When $\alpha \neq 1$, it is necessary to employ a preliminary transformation which is suggested by the work in that part of Barnes' memoir on integral functions which deals with the generalised function of Mittag-Leffler:

$$E_\alpha(x;\theta,\beta)=\sum_{n=0}^{\infty}\frac{x^n}{\Gamma(1+\alpha n).(n+\theta)^\beta}.$$

The transformation is derived by considering the contour integral

$$I_4=-\frac{1}{2\pi i}\int_D \frac{\pi\sin\{(q-\alpha)\pi s\}}{\sin(\alpha\pi s).\sin(\pi s)}\cdot\frac{\phi(s).e^{gs}}{\Gamma(\alpha s+1)}.x^s ds \quad\ldots\ldots\ldots\ldots(7);$$

the contour of integration is parallel to* the line $R(\alpha s)=h_1$, and on the right of this line; let $h$ and $j$ be the integers such that the points $s=j$, $s=h/\alpha$, $s=-\Theta/\alpha$ lie on the left of $D$ while the points $s=j+1$, $s=(h+1)/\alpha$ lie on the right of $D$; also let $q$ be an odd positive integer $(=2p+1)$; we defer, for the moment, the choice of the numbers $q$ and $p$.

The integral $I_4$ converges if

$$|\sin\gamma\log|z|-\cos\gamma\arg z|<\tfrac{1}{2}\pi A+\pi\cos\gamma-\pi|(q\cos\gamma-A)| \quad\ldots\ldots(7\,a),$$

where $$z=e^g x,\quad \alpha=Ae^{i\gamma},$$

the numbers $A$ and $\gamma$ being real and $|\gamma|<\frac{1}{2}\pi$.

[We notice that $h+1>-R(\Theta)$ so that $h+1-\mu>0$.]

Now if (7 *a*) is satisfied we may shew that $I_4$ is $O(|x|^L)$ where $L$ is the value of $R(\alpha s)/R(\alpha)$ on the contour $D$. Further, we may shew, by Cauchy's theorem, that $I_4$ is equal to minus the sum of the residues of the integrand at those poles which lie on the right of $D$.

We thus get

$$I_4=\sum_{n=j+1}^{\infty}\frac{\phi(n)e^{gn}}{\Gamma(\alpha n+1)}x^n-\sum_{n=h+1}^{\infty}\frac{\sin(\pi q n/\alpha)}{\sin(\pi n/\alpha)}\frac{\phi(n/\alpha)}{\Gamma(n+1)}(xe^g)^{n/\alpha}.$$

Remembering that $q=2p+1$, we get

$$\sum_{n=j+1}^{\infty}\frac{\phi(n).e^{gn}}{\Gamma(\alpha n+1)}x^n=I_4+\sum_{n=h+1}^{\infty}\sum_{t=-p}^{p}\exp(2\pi int/\alpha).\frac{\phi(n/\alpha)}{\Gamma(n+1)}(xe^g)^{n/\alpha}\ldots\ldots\ldots(8).$$

We are thus led to consider the behaviour of the integral function

$$F(y)=\sum_{n=h+1}^{\infty}y^n\frac{\phi(n/\alpha)}{\Gamma(n+1)}\quad\ldots\ldots\ldots\ldots\ldots\ldots(9),$$

where $$y=(xe^g)^{1/\alpha}\exp(2\pi it/\alpha).$$

But, for the values of $n$ involved in this summation, we may expand $\phi(n/\alpha)$ into a convergent series of inverse factorials (by the result of § 2), so that

$$\phi(n/\alpha)=\frac{1}{(n+\Theta)^\beta}\left[a_0+\frac{a_1}{n+1}+\frac{a_2}{(n+1)(n+2)}+\ldots\right].$$

* We take the contour to lie on the right of the line $R(\alpha s)=\gamma'$; $\gamma'$ being defined as in § 2.

(On comparing both sides of this equation as $n \to \infty$, we see that the first $\beta' - \beta$ of the coefficients $a_0, a_1, \dots$ vanish.)

From this expansion, we may shew that $F(y)$ can be expressed as the sum of a series of functions of the type $G_\beta(y; \theta, k)$; for we have

$$F(y) = \sum_{m=0}^{k} \sum_{n=h+1}^{\infty} \frac{a_m y^n}{(n+\Theta)^\beta \Gamma(n+m+1)} + \sum_{n=h+1}^{\infty} \frac{b_{k,n} y^n}{(n+\Theta)^\beta \Gamma(n+1)},$$

where
$$b_{k,n} = \frac{a_{k+1}}{(n+1)(n+2)\dots(n+k+1)} + \frac{a_{k+2}}{(n+1)(n+2)\dots(n+k+2)} + \dots.$$

Also, since $n \geqslant h+1$,

$$|b_{k,n}| < (h+1)! \left[\frac{|a_{k+1}|}{(h+k+2)!} + \frac{|a_{k+2}|}{(h+k+3)!} + \dots\right],$$

and, since $|a_{k+1}| < Kk!\,k^\mu \{\log(k+2)\}^B$, we easily find that, *qua* function of $k$ and $n$,

$$|b_{k,n}| < (h+1)!\,K \left[\frac{\{\log(k+2)\}^B}{k^{h+2-\mu}} + \frac{\{\log(k+3)\}^B}{(k+1)^{h+2-\mu}} + \dots\right]$$

$$= O\left[\int_k^\infty \frac{\{\log(t+2)\}^B}{t^{h+2-\mu}}\,dt\right]$$

$$= O\left[\frac{\log(k+2)^B}{k^{h+1-\mu}}\right]$$

$$\to 0, \text{ as } k \to \infty \text{ since } h+1-\mu > 0.$$

Further, if $b_{(k)}$ be the greatest value of $|b_{k,n}|$ for $n \geqslant h+1$, we have

$$\left|\sum_{n=h+1}^{\infty} \frac{b_{k,n} y^n}{(n+\Theta)^\beta \Gamma(n+1)}\right| \leqslant b_{(k)} \sum_{n=h+1}^{\infty} \frac{|y|^n}{|(n+\Theta)^\beta| \,.\, n!}.$$

Since $R(n+\Theta) > 0$ throughout the summation, $\sum_{n=h+1}^{\infty} \frac{|y|^n}{|(n+\Theta)^\beta| \,.\, n!}$ is finite for any assigned value of $y$. And consequently

$$F(y) = \sum_{m=0}^{k} a_m y^{h+1} G_\beta(y;\ \Theta+h+1,\ m+h+1) + J_k,$$

where $|J_k| \leqslant b_{(k)} \sum_{n=h+1}^{\infty} \frac{|y|^n}{|(n+\Theta)^\beta| \,.\, n!} \to 0$, as $k \to \infty$ when $y$ is assigned.

That is to say, for every finite value of $y$,

$$F(y) = \sum_{m=0}^{\infty} a_m y^{h+1} G_\beta(y;\ \Theta+h+1,\ m+h+1).$$

6. *We shall now shew that when* $|\arg y| < \frac{1}{2}\pi$, $F(y)$ *possesses the asymptotic expansion*

$$F(y) = \frac{e^y}{y^\beta}\left[\sum_{n=0}^{\nu} \frac{S_n \,.\, \Gamma(1-\beta)}{y^n} + o(|y|^{-\nu})\right] \dots\dots\dots\dots (10)$$

where $S_n = \sum_{m=0}^{n} \frac{a_m \,.\, {}_mc_{n-m}}{\Gamma(1-\beta-n+m)}$; the coefficients ${}_mc_n$ are given by the expansion

$$[\log(1-x)]^{\beta-1}(1-x)^{\theta-m-1} \equiv (-x)^{\beta-1} \sum_{n=0}^{\infty} {}_mc_n (-x)^n,$$

and $\theta = \Theta + h + 1$ while $\nu$ is any fixed integer.

*We shall also shew that when* $|\arg(-y)| \leqslant \frac{1}{2}\pi$, $F(y)$ *is* $O(|y|^{\lambda})$ *where* $\lambda$ *is determinate.*

We have

$$F(y) = \sum_{m=0}^{\nu} a_m y^{h+1} G_\beta(y;\ \Theta + h + 1,\ m + h + 1) + \sum_{m=\nu+1}^{\infty} a_m y^{h+1} G_\beta(y;\ \Theta + h + 1,\ m + h + 1).$$

On substituting the result (6c) into the first series in this equation, we get at once when $|\arg y| \leqslant \pi$ and $|\arg(-y)| \leqslant \pi$,

$$F(y) = \frac{e^y}{y^\beta}\left[\sum_{n=0}^{\nu} \frac{S_n \,.\, \Gamma(1-\beta)}{y^n} + o(|y|^{-\nu})\right] + O(|y|^{h}) + O(|y^{-\Theta}\{\log(-y)\}^{\beta-1}|)$$

$$+ \sum_{m=\nu+1}^{\infty} a_m y^{h+1} G_\beta(y;\ \Theta + h + 1,\ m + h + 1).$$

Making use of the formulae (6) and (6a) we find, on remembering that $R(\Theta + h + 1) > 0$, that

$$\left|\sum_{m=\nu+1}^{\infty} a_m y^{h+1} G_\beta(y;\ \Theta + h + 1,\ m + h + 1)\right|$$

$$< \left|\frac{K\, e^y y^{h+1}}{y^{N+l+\beta-1}}\right| \sum_{m=\nu+1}^{\infty} \left|\frac{a_m (m+h+1)^N}{\Gamma(m+h+2)}\right| + K\,|y^{2-\beta+h}| \sum_{m=\nu+1}^{\infty} \left|\frac{a_m (m+h+1)}{\Gamma(m+h+2)}\right|$$

$$+ K\,|y^{h+1}| \sum_{m=\nu+1}^{\infty} \left|\frac{a_m (m+h+1)^{R(\Theta+h+1)}}{\Gamma(m+h+2)\,\{\log(m+h+1)\}^{1-R(\beta)}}\right|,$$

where $N$ is the integer such that $N < R(\Theta + h + 1) < N + 1$, and $l$ is any *fixed* positive integer, $l$ and $\nu$ being so chosen that $\nu > N + l$.

Remembering that, for the values of $m$ under consideration, $a_m$, *quà* function of $m$, is such that $a_m = O(\Gamma(m)\{\log(m+1)\}^B m^\mu)$, we see without difficulty that

$$\frac{a_m (m+h+1)^N}{\Gamma(m+h+2)} = O\left(\frac{\{\log m\}^B}{m^{h+2-N-\mu}}\right),$$

and*

$$\frac{a_m (m+h+1)^{R(\Theta+h+1)}}{\Gamma(m+h+2)\,\{\log(m+h+1)\}^{1-R(\beta)}} = O\left(\frac{1}{m\,\{\log m\}^{1-R(\beta)-B}}\right).$$

Now $h + 1 - N - \mu > 0$ since $R(\Theta + h + 1) > N$, and $R(\beta) + B < 0$; and hence the series $\sum_{m=2}^{\infty} \frac{\{\log m\}^B}{m^{h+2-N-\mu}}$, $\sum_{m=2}^{\infty} \frac{1}{m\,\{\log m\}^{1-R(\beta)-B}}$ are absolutely convergent.

Hence it is evident that

$$\left|\sum_{m=\nu+1}^{\infty} a_m y^{h+1} G_\beta(y;\ \Theta + h + 1,\ m + h + 1)\right| < \left|\frac{K\, e^y y^{h+1}}{y^{N+l+\beta-1}}\right| + K\,|y^{2-\beta+h}| + K\,|y^{h+1}|.$$

Consequently, when $|\arg y| \leqslant \pi$,

$$F(y) = \frac{e^y}{y^\beta}\left[\sum_{n=0}^{\nu} \frac{S_n \,.\, \Gamma(1-\beta)}{y^n} + o(|y|^{-\nu}) + O(|y|^{h+2-l-N})\right] + O(|y|^{\lambda}),$$

where $\lambda$ is the greatest of the numbers $h+1$, $1 - R(\Theta)$, $2 - R(\beta) + h$.

This result may be written in the form:

$$F(y) = \frac{e^y}{y^\beta}\left[\sum_{n=0}^{N+l-h-1} \frac{S_n \,.\, \Gamma(1-\beta)}{y^n} + o(|y|^{h+1-l-N})\right] + O(|y|^{\lambda}),$$

* $[\mu = -R(\Theta)]$.

and $\nu$ may be chosen in such a way that $N+l-h-1$ is any fixed positive integer; for $\nu$ is any fixed integer and $l$ is subject only to the condition that $\nu > N+l$.

Writing $\nu$ in place of $N+l-h-1$, we see that this result may be written in the form:

$$F(y) = \frac{e^y}{y^\beta}\left[\sum_{n=0}^{\nu} \frac{S_n \,.\, \Gamma(1-\beta)}{y^n} + o(|y^{-\nu}|)\right] + O(|y^\lambda|),$$

when $|\arg y| \leqslant \pi$.

This is, effectively, the result stated at the beginning of this section.

7. If we substitute the asymptotic formula of § 6 for $F(y)$ in the equation (8) we see that if one or more of the numbers

$$|\arg z^{1/a} + \arg\{\exp(2\pi i\mu/\alpha)\}| \qquad (\mu = -p, 1-p, 2-p, \ldots, +p)$$

is less than $\frac{1}{2}\pi$, and if none are greater than $\frac{3}{2}\pi$, the asymptotic expansion of $f(x)$ is given by the formula

$$f(x) = \sum_{l=-p}^{p}\left[\frac{\exp(x^{1/a}e^{g/a+2\pi i\mu/a})}{(x^{1/a}e^{g/a+2\pi i\mu/a})^\beta} \times \left\{\sum_{n=0}^{\nu} \frac{S_n \,.\, \Gamma(1-\beta)}{(x^{1/a}e^{g/a+2\pi i\mu/a})^n} + o(|x^{-\nu/a}|)\right\}\right] \quad \ldots\ldots(11),$$

where $\nu$ is any fixed positive integer.

8. In order to appreciate clearly the range of validity of the asymptotic expansion (11), we have to consider the inequality (7 $a$), viz.

$$|\{\sin\gamma \log|z| - \cos\gamma \arg z\}| < \tfrac{1}{2}\pi A + \pi\cos\gamma - \pi|(q\cos\gamma - A)|,$$

in some detail.

(A) In the first instance let us suppose that

$$A < 2\cos\gamma;$$

this inequality may also be written $|\alpha - 1| < 1$.

When $z$ lies in the portion of the plane* defined by the inequalities

$$-\tfrac{1}{2}\pi A < \sin\gamma\log|z| - \cos\gamma\arg z < \tfrac{1}{2}\pi A \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots(12),$$

then it is easy to see that $R(x^{1/a}e^{g/a}) > 0$ and the inequality (7 $a$) is satisfied if $q = 1$.

In other words, if $|\alpha - 1| < 1$ and $|\sin\gamma\log|z| - \cos\gamma\arg z| < \frac{1}{2}\pi A$, then $f(x)$ possesses the asymptotic expansion

$$f(x) = \frac{\exp(x^{1/a}e^{g/a})}{x^{\beta/a}e^{g\beta/a}}\left\{\sum_{n=0}^{\nu}\frac{S_n \,.\, \Gamma(1-\beta)}{x^{n/a}e^{gn/a}} + o(|x^{-\nu/a}|)\right\} \quad \ldots\ldots\ldots\ldots\ldots\ldots(13).$$

(B) If, however, $x$ does not lie in the region of the plane which we have been considering in (A), we may choose $\arg(-z)$ so that

$$-(\pi\cos\gamma - \tfrac{1}{2}\pi A) < \sin\gamma\log|z| - \cos\gamma\arg(-z) < \pi\cos\gamma - \tfrac{1}{2}\pi A \quad \ldots\ldots\ldots(14).$$

And, when $z$ is such that (14) is satisfied, we may shew that

$$f(x) - \sum_{n=0}^{j}\frac{\phi(n)e^{gn}x^n}{\Gamma(\alpha n+1)} = \frac{1}{2\pi i}\int_D \frac{\pi \,.\, \phi(s)(-z)^s}{\sin(\pi s)\,\Gamma(\alpha s+1)}\,ds \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots(15).$$

If the subject of integration be a uniform function of $s$ and if it tends to zero as $s \to \infty$ when $-h_2 \leqslant R(\alpha s) \leqslant h_1$ for every fixed value of $h_2$, a knowledge of the poles of $\phi(s)$ on the left

* That is to say, $x$ is confined to a region of the plane between two equiangular spirals.

of the line $R(\alpha s) = h_1$ will yield the asymptotic expansion of $f(x)$ in the usual manner; for we have

$$f(x) - \sum_{n=0}^{j} \frac{\phi(n) e^{gn} x^n}{\Gamma(\alpha n + 1)} = \frac{1}{2\pi i} \int_E \frac{\pi . \phi(s) (-z)^s \, ds}{\sin(\pi s) \Gamma(\alpha s + 1)} + \text{the sum of the residues of } \frac{\pi . \phi(s) (-z)^s}{\sin(\pi s) \Gamma(\alpha s + 1)}$$

at those poles which lie between $D$ and $E$; $E$ being a contour parallel to $D$ and on the left of it.

If on $E$, $R(\alpha s)/R(\alpha) = -h_2$, the integral along $E$ is $O(|z|^{-h_2})$. If, however, $h_2$ could not exceed some definite value, we should not get a complete asymptotic expansion, but we should obtain an asymptotic formula containing a definite number of terms, valid over the region specified by (14).

In the immediate vicinity of the curves

$$\sin\gamma \log|z| - \cos\gamma \arg z = \pm \tfrac{1}{2}\pi A,$$

the only result concerning $f(x)$ which we can obtain is an equation of the form

$$f(x) = O(|x|^{\lambda_1}),$$

where $\lambda_1$ is determinate; this follows without difficulty from (10).

(C) Let $A = 2\cos\gamma$, i.e. $|\alpha - 1| = 1$. ($A \neq 0$.)

By taking $q = 1$, we get the expansion (13) as before, valid over the region

$$-\pi\cos\gamma < \sin\gamma \log|z| - \cos\gamma \arg z < \pi\cos\gamma.$$

That is to say, we have obtained the asymptotic expansion of $f(x)$ over the whole plane with the exception of the neighbourhood of the spiral

$$\arg z - \tan\gamma \log|z| = \pm \pi.$$

Near this spiral, we only know that $f(x) = O(x^{\lambda_1})$; and we get no more information by taking $q = 3, 5, \ldots$.

(D) Lastly let $A > 2\cos\gamma$, i.e. $|\alpha - 1| > 1$.

Let us choose $q$ to be such an odd integer that

$$A - \cos\gamma \leqslant q\cos\gamma < A + \cos\gamma \quad \text{.................................(16).}$$

Whatever be the position of $x$ in the plane, we can always choose $\arg z$ so that the inequalities

$$-\tfrac{1}{2}\pi A < \sin\gamma \log|z| - \cos\gamma \arg z < \tfrac{1}{2}\pi A \quad \text{.......................(16}a\text{)}$$

are satisfied; so that with this determination of $\arg z$ and $\arg x$ we have

$$R(x^{1/\alpha} e^{g/\alpha}) > 0, \quad |\arg(x^{1/\alpha} e^{g/\alpha})| < \tfrac{1}{2}\pi;$$

and then ($7a$) is satisfied since $\cos\gamma - |(q\cos\gamma - A)| \geqslant 0$.

Further, we have to ensure that the arguments of all the expressions $x^{1/\alpha} e^{g/\alpha} e^{2\pi i\mu/\alpha}$ lie between $\pm \tfrac{3}{2}\pi$. But this is certainly the case, for

$$|\arg(x^{1/\alpha} e^{g/\alpha} e^{2\pi i\mu/\alpha})| < \tfrac{1}{2}\pi + \arg(e^{2\pi i p/\alpha})$$
$$< \tfrac{1}{2}\pi + \pi p A^{-1}\cos\gamma < \tfrac{1}{2}\pi\{1 + (q-1)A^{-1}\cos\gamma\} < \pi.$$

Thus, when $|\alpha - 1| > 1$, the asymptotic expansion of $f(x)$, valid over the whole plane, is given by

$$f(x) = \sum_{\mu=-p}^{p} \left[ \frac{\exp(x^{1/\alpha} e^{g/\alpha + 2\pi i\mu/\alpha})}{x^{\beta/\alpha} e^{g\beta/\alpha + 2\pi i\mu\beta/\alpha}} \times \left\{ \sum_{n=0}^{\nu} \frac{S_n . \Gamma(1-\beta)}{(x^{1/\alpha} e^{g/\alpha + 2\pi i\mu/\alpha})^n} + o(|x^{-\nu/\alpha}|) \right\} \right],$$

where $\arg x$ is determined by the inequality

$$[\sin\gamma\{\log|x| + R(g)\} - \cos\gamma\{\arg x + I(g)\}] < \tfrac{1}{2}\pi A,$$

$q = 2p+1$, $q$ is defined by (16) and the coefficients $S_n$ are those defined in connection with equation (10).

Since the first $\beta' - \beta$ of the coefficients $a_0, a_1, \ldots$ vanish, so also do the first $\beta' - \beta$ of the coefficients $S_0, S_1, \ldots$; it will, therefore, be possible to modify the above expansion slightly when $\beta' - \beta$ is a positive integer and not zero; we shall not carry out this modification, as a more convenient form of the expansion will be obtained in Part IV.

## PART III. *Asymptotic formulae required when $M < 1$.*

9. The fundamental function to be investigated in place of $E_k(x)$, when $M<1$, is the integral function

$$E_k(x;\, M, c) = \sum_{n=0}^{\infty} \frac{x^n}{n!\,\{M(n+c)+1\}\{M(n+c)+2\}\ldots\{M(n+c)+k\}}.$$

The asymptotic formulae are required for this function when $c$ is an assigned positive number, $k$ is a (large) integer and $x$ has any value.

It is evident that $$E_k(x;\, M, c) = \frac{1}{\Gamma(k)} \sum_{n=0}^{\infty} \frac{x^n}{n!} \frac{\Gamma(Mn+Mc+1)\,\Gamma(k)}{\Gamma(Mn+Mc+k+1)}.$$

Now $\sum_{n=0}^{\infty} \frac{x^n}{n!} t^{M(n+c)}(1-t)^{k-1}$ is uniformly convergent and each term of the series is continuous (when $k \geqslant 1$) for the range $0 \leqslant t \leqslant 1$ for all values of $x$; consequently* we may integrate term-by-term and we get

$$E_k(x;\, M, c) = \frac{1}{\Gamma(k)} \int_0^1 (1-t)^{k-1} t^{Mc} \exp(xt^M)\, dt.$$

If $R(x) \geqslant 0$, $|\exp(xt^M)| \leqslant |\exp x|$ on the path of integration.

Hence, when $R(x) \geqslant 0$,

$$|E_k(x;\, M, c)| \leqslant \frac{|e^x|}{\Gamma(k)} \int_0^1 (1-t)^{k-1} t^{Mc}\, dt,$$

i.e. $$|E_k(x;\, M, c)| \leqslant \frac{|e^x|\,\Gamma(Mc+1)}{\Gamma(Mc+k+1)} \quad \text{.......................................(17 } a\text{)}.$$

In like manner, when $R(x) \leqslant 0$, $|\exp(xt^M)| \leqslant 1$; so that, when $R(x) \leqslant 0$,

$$|E_k(x;\, M, c)| \leqslant \frac{\Gamma(Mc+1)}{\Gamma(Mc+k+1)} \quad \text{.......................................(17 } b\text{)}.$$

Further, we have $|e^{-x} E_k(x;\, M, c)| \leqslant \frac{1}{\Gamma(k)} \int_0^1 (1-t)^{k-1} t^{Mc} \exp\{R(x)(t^M-1)\}\, dt.$

Now, if $R(x)>0$, $(1-t^M)^l \exp\{R(x)\,.\,(t^M-1)\}$, *qua* function of $t$, has one (and only one) maximum when $0 \leqslant t \leqslant 1$, viz. when $1 - t^M = \frac{l}{R(x)}$, provided that $R(x) \geqslant l$.

Consequently if $R(x) \geqslant l$ and $k \geqslant l+1$ we have

$$(1-t^M)^l \exp\{R(x)\,.\,(t^M-1)\} \leqslant \left(\frac{le^{-1}}{R(x)}\right)^l \text{ when } 0 \leqslant t \leqslant 1.$$

So that if $l$ be *fixed*,

$$|e^{-x} E_k(x;\, M, c)| < \frac{K}{\Gamma(k)\,\{R(x)\}^l} \int_0^1 (1-t)^{k-1} t^{Mc} (1-t^M)^{-l}\, dt.$$

But, when $M \leqslant 1$, $(1-t)/(1-t^M) \leqslant M^{-1}$ when $0 \leqslant t \leqslant 1$,

so that $$|e^{-x} E_k(x;\, M, c)| < \frac{M^{-l} K}{\Gamma(k)\,\{R(x)\}^l} \int_0^1 t^{Mc}(1-t)^{k-l-1}\, dt < \frac{K}{\Gamma(k)\,\{R(x)\}^l} \frac{\Gamma(k-l)\,\Gamma(Mc+1)}{\Gamma(Mc+k-l+1)}.$$

* See Bromwich, *Theory of Infinite Series*, p. 116.

But if $k \geqslant l+1$, $\dfrac{Mc+k-l+r}{k-l+r-1} < \dfrac{Mc+r+1}{r}$ where $r=1, 2, \dots l$, so that

$$\frac{(Mc+k-l+1)(Mc+k-l+2)\dots(Mc+k)}{(k-l)(k-l+1)\dots(k-1)} < \frac{(Mc+2)(Mc+3)\dots(Mc+l+1)}{1 \,.\, 2 \dots l}.$$

Consequently, if $l$ be an assigned integer $k \geqslant l+1$ and $R(x) \geqslant l$, then

$$E_k(x;\ M, c) < \frac{K}{\{R(x)\}^l} \cdot \frac{\Gamma(k-l)}{\Gamma(k)} \cdot \frac{\Gamma(Mc+k+1)}{\Gamma(Mc+k-l+1)} \cdot \frac{\Gamma(Mc+l+1)}{\Gamma(Mc+k+1)} e^{x},$$

i.e.
$$|E_k(x;\ M, c)| < \frac{K e^{|x|}}{\{R(x)\}^l \,\Gamma(Mc+k+1)} \quad \text{..........................................(17 } c\text{)}.$$

And, by (17 $a$), this inequality is true when $0 \leqslant R(x) \leqslant l$.

Lastly, we get on integrating by parts, when $R(x) \leqslant 0$,

$$E_k(x;\ M, c) = \frac{1}{\Gamma(k)}\left[\frac{1}{Mx}(1-t)^{k-1} t^{Mc-M-1} \exp(xt^M)\right]_0^1 - \frac{1}{\Gamma(k)}\int_0^1 x^{-1} M^{-1} \exp(xt^M) \,.\, \frac{d}{dt}\{(1-t)^{k-1} t^{Mc-M+1}\} \,.\, dt,$$

i.e. if $k>1$, $E_k(x;\ M, c) = -\dfrac{1}{\Gamma(k)}\displaystyle\int_0^1 \frac{\exp(xt^M)}{Mx}\{-(k-1)(1-t)^{k-2} t^{Mc-M+1} + (Mc-M+1)(1-t)^{k-1} t^{Mc-M}\}\, dt.$

Since $|\exp(xt^M)| \leqslant 1$ when $R(x) \leqslant 0$, we see that

$$|E_k(x;\ M, c)| \leqslant \frac{1}{M(x)\,\Gamma(k)}\int_0^1 \{(k-1)(1-t)^{k-2} t^{Mc-M+1} + (Mc-M+1) \,.\, (1-t)^{k-1} t^{Mc-M}\}\, dt$$
$$\leqslant \frac{2\Gamma(Mc-M+2)}{M|x|\,\Gamma(Mc-M+k+1)}.$$

The reader will easily see, on making the necessary modifications in the work, that this is true when $k=1$.

Hence when $R(x) \leqslant 0$, we have, in addition to (17 $b$),

$$|E_k(x;\ M, c)| < \frac{2\Gamma(Mc-M+2)}{M|x|\,\Gamma(Mc-M+k+1)} \quad \text{....................................(17 } d\text{)}.$$

The results (17) are all true when $M=1$.

10. We now have to discuss the asymptotic behaviour of the integral function

$$G_\beta(x;\ \theta, k;\ M, c) = \sum_{n=0}^{\infty} \frac{x^n}{n!\,(n+\theta)^\beta \{M(n+c)+1\}\{M(n+c)+2\}\dots\{M(n+c)+k\}},$$

when $|x|$ is large, $k$ is a (large) integer, $0 < R(\theta) < 1$, $R(\beta) < 0$, $c = R(\theta)$ and $M \leqslant 1$.

As in § 4, we may shew that

$$G_\beta(x;\ \theta, k;\ M, c) = \frac{e^{-\pi i\beta}\,\Gamma(1-\beta)}{2\pi i x}\int\left[-\log\left(1-\frac{y}{x}\right)\right]^{\beta-1}\left(1-\frac{y}{x}\right)^{\theta-1} E_k(x-y;\ M, c)\, dy,$$

round the contour marked with double arrows in Fig. 1; and, as in § 4, we may deform the contour into that marked with single arrows in Fig. 1, so that, preserving the notation of § 4, we have

$$G_\beta(x;\ \theta, k;\ M, c) = \frac{e^{-\pi i\beta}\,\Gamma(1-\beta)}{2\pi i x}\int_C\left[-\log\left(1-\frac{y}{x}\right)\right]^{\beta-1}\left(1-\frac{y}{x}\right)^{\theta-1} E_k(x-y;\ M, c) \,.\, dy$$
$$-\frac{1}{x\,\Gamma(\beta)}\int_{PQ}\left[-\log\left(1-\frac{y}{x}\right)\right]^{\beta-1}\left(1-\frac{y}{x}\right)^{\theta-1} E_k(x-y;\ M, c) \,.\, dy$$
$$-\frac{1}{x\,\Gamma(\beta)}\int_{OP'}\left[-\log\left(-\frac{\eta}{x}\right)\right]^{\beta-1}\left(-\frac{\eta}{x}\right)^{\theta-1} E_k(-\eta;\ M, c) \,.\, d\eta \quad \text{......(18)}$$
$$= J_1 + J_2 + J_3 \quad \text{.........................................................................(18}a\text{)}.$$

On $PQ$, as in § 4, $\left|\left[-\log\left(1-\frac{y}{x}\right)\right]^{\beta-1}\left(1-\frac{y}{x}\right)^{\theta-1}\right| < K$, and since, on $PQ$, $R(x-y)<0$, $|E_k(x-y;\ M, c)| < K\{\Gamma(Mc+k+1)\}^{-1}$; while the length $PQ$ is less than or equal to $2|x|$.

Hence
$$|J_2| < \frac{1}{|x\,\Gamma(\beta)|}\int_{PQ}\frac{K}{\Gamma(Mc+k+1)}|dy|,$$

i.e.
$$|J_2| < \frac{K}{\Gamma(Mc+k+1)} \quad\text{.................................(19}a\text{).}$$

We have now to estimate the value of $|J_3|$; putting $\eta = rxe^{i\omega}$ where $r$ is real and positive, we get

$$|J_3| < \frac{1}{|x\,\Gamma(\beta)|}\int_0^\rho |(-\log r - i\omega)^{\beta-1}|\,.\,|(re^{i\omega})^{\theta-1}|\,.\,|E_k(-\eta;\ M, c)|\,.\,|x|\,dr,$$

where $\rho < 6k\exp|x|$; whence we deduce that

$$|J_3| < K\int_0^\rho \frac{r^{c-1}\,dr}{\{|\log r| + |\omega|\}^{1-R(\beta)}}\,.\,|E_k(-\eta;\ M, c)|.$$

We now divide the range of integration into three parts, viz.

(i) from $r=0$ to $r=k_2$ where $k_2 = k^{\frac{1}{2}M}$,

(ii) from $r=k_2$ to $r=k^M$,

(iii) from $r=k^M$ to $r=\rho$;

we observe that:

on (i), $\{|\log r| + |\omega|\}^{R(\beta)-1} < K$,

on (ii), $\{|\log r| + |\omega|\}^{R(\beta)-1} < K\{\log k_2\}^{R(\beta)-1}$,

on (iii), $\{|\log r| + |\omega|\}^{R(\beta)-1} < K\{\log k\}^{R(\beta)-1}$;

also in (i) and (ii) we use the inequality $|E_k(-\eta)| < K\{\Gamma(Mc+k+1)\}^{-1}$ derived from (17$b$); and in (iii) we use the inequality $|E_k(-\eta)| < K\{|\eta|(Mc-M+k+1)\}^{-1}$ derived from (17$d$); and we deduce that

$$\begin{aligned}|J_3| &< K\int_{(i)}\frac{r^{c-1}}{\Gamma(Mc+k+1)}dr + K\int_{(ii)}\frac{r^{c-1}}{\Gamma(Mc+k+1)}dr\,\{\log k_2\}^{R(\beta)-1}\\ &\qquad + K\int_{(iii)}\frac{r^{c-2}}{|x|\,\Gamma(Mc-M+k+1)}dr\,\{\log k\}^{R(\beta)-1}\\ &< K\left[\frac{k_2^c + k^{Mc}\{\log k_2\}^{R(\beta)-1}}{\Gamma(Mc+k+1)} + \frac{k^{M(c-1)}\{\log k\}^{R(\beta)-1}}{|x|\,\Gamma(Mc-M+k+1)}\right],\end{aligned}$$

since $0<c<1$; on substituting for $k_2$ and making use of Stirling's formula, we may write this result in the form

$$|J_3| < \frac{K}{\Gamma(k+1)\{\log k\}^{1-R(\beta)}} \quad\text{.................................(19}b\text{).}$$

Next we have to estimate the value of the integral

$$\int_C\left[-\log\left(1-\frac{y}{x}\right)\right]^{\beta-1}\left(1-\frac{y}{x}\right)^{\theta-1}E_k(x-y;\ M, c)\,dy.$$

We divide the path of integration into two portions, the first portion being such that on it $|y| \leqslant 2|x|$, the second being such that on it $|y| \geqslant 2|x|$; we call the integrals, along these portions, $J_1'$ and $J_1''$ respectively.

As in investigating $I_1'$ in § 4, we may shew that

$$|J_1'| < \frac{K}{|x^{\beta-1}|} \int |E_k(x-y;\ M, c)\, dy| \quad \text{..........................(19}c\text{)}.$$

We now have two cases to consider, according as $R(x) \leqslant 0$ or $R(x) \geqslant 0$.

(A) If $R(x) \leqslant 0$ then $R(x-y) \leqslant 1$ throughout the contour and the length of the contour is less than $k|x|$; so that by substituting for $E_k(x-y;\ M, c)$ from (17$a$) and (17$b$), we get

$$|J_1'| < \frac{K|x|}{|x^{\beta-1}|\, \Gamma(Mc+k+1)},$$

i.e.

$$|J_1'| < \frac{K}{|x^{\beta-2}|\, \Gamma(Mc+k+1)} \quad \text{..............................(19}d\text{)},$$

when $R(x) \leqslant 0$.

(B) If $R(x) \geqslant 0$, let $\arg x = \Omega$ where $|\Omega| \leqslant \frac{1}{2}\pi$. If $\Omega$ is not actually equal to $\pm\frac{1}{2}\pi$, we can divide the path of integration into two parts, on one of which $R(x-y) \geqslant \frac{1}{2}R(x)$, and on the other $R(x-y) \leqslant \frac{1}{2}R(x)$; and the length of each of these parts is less than $K|x|$.

On the first part we have from (17$c$), if $k \geqslant l+1$,

$$|E_k(x-y;\ M, c)| < \frac{K|e^{x-y}|}{\{R(x-y)\}^l\, \Gamma(Mc+k+1)},$$

i.e., since $R(x-y) \geqslant \frac{1}{2}R(x)$,

$$|E_k(x-y;\ M, c)| < \frac{K(\sec\Omega)^l\, |e^x|\,.\,|e^{-y}|}{|x^l|\, \Gamma(Mc+k+1)}, \text{ and } |e^{-y}| < e.$$

On the second part, by (17$a$),

$$|E_k(x-y;\ M, c)| < \frac{K|e^{x-y}|}{\Gamma(Mc+k+1)}$$

$$< \frac{K|e^{\frac{1}{2}x}|}{\Gamma(Mc+k+1)}.$$

Hence, by (19$c$), we see that, when $\arg x = \Omega$, $R(x) > 0$ and $k \geqslant l+1$,

$$|J_1'| < \frac{K(\sec\Omega)^l\, |e^x|}{|x^{l+\beta-2}|\, \Gamma(Mc+k+1)} + \frac{K|e^{\frac{1}{2}x}|}{|x^{\beta-2}|\, \Gamma(Mc+k+1)} \quad \text{..................(19}e\text{)}.$$

A modified form of this result is desirable when $\Omega$ is nearly equal to $\pm\frac{1}{2}\pi$. To obtain it, we notice that, on the path of integration, $R(y) \geqslant -1$, and hence by (17$a$) and (17$b$), when $R(x) \geqslant 0$,

$$|E_k(x-y;\ M, c)| < \frac{K|e^x|}{\Gamma(Mc+k+1)}.$$

From this result we derive, by (19$c$), the following inequality when $R(x) \geqslant 0$:

$$|J_1'| < \frac{K|e^x|}{|x^{\beta-2}|\, \Gamma(Mc+k+1)} \quad \text{..............................(19}f\text{)}.$$

Lastly, we have to estimate the value of the integral

$$J_1'' = \int \left[-\log\left(1-\frac{y}{x}\right)\right]^{\beta-1} \left(1-\frac{y}{x}\right)^{\theta-1} E_k(x-y;\ M, c)\, dy,$$

taken along that portion of the contour $C$ for which $|y| \geqslant 2|x|$; on this portion of the contour the formulae (17 $b$) and (17 $d$) are applicable.

By putting $1 - y/x = re^{i\omega}$ ($r$ real) and dividing the path of integration into three portions, viz. (i) from the point where $|y| = 2|x|$ to the point where $|1 - y/x| = k^{\frac{1}{2}M}$; (ii) from this point to the point where $|1 - y/x| = k^{M}$; (iii) from this point to the point where $y = OQ$, and on making use of (17 $b$) in (i) and (ii) and of (17 $d$) in (iii) we find, in just the same way as that in which we obtained (5 $d$), that

$$|J_1''| < \frac{K|x|}{\Gamma(k+1)\{\log k\}^{1-R(\beta)}} \quad \text{.................................(19}g\text{)}.$$

Collecting the results numbered (19) we find that the asymptotic inequality for $G_\beta(x;\ \theta, k;\ M, c)$ may be written in the following form:

Let $l$ be a fixed integer, $k$ a (large) integer such that $k \geqslant l + 1$, and let $|x| > 2$, $\arg x = \Omega$.

Then

$$|G_\beta(x;\ \theta, k;\ M, c)| < \frac{KU}{|x^{\beta-1}|\,\Gamma(Mc+k+1)} + \frac{K|x^{1-\beta}|}{\Gamma(Mc+k+1)} + \frac{K}{\Gamma(k+1)\{\log k\}^{1-R(\beta)}} \quad \text{...(20)},$$

where $U$ is defined by the following inequalities:

(i) when $R(x) \leqslant 0$, $U = 1$ ..............................................................(20 $a$),

(ii) when $R(x) \geqslant 0$, *both* the inequalities,

$$U < \frac{(\sec\Omega)^l}{|x^l|}|e^x| + |e^{\frac{1}{2}x}| \quad \text{...................................(20}b\text{)},$$

$$U < |e^x| \quad \text{.......................................................(20}c\text{)},$$

are true.

## PART IV. *The asymptotic expansion of $f(x)$ when $M < 1$.*

11. The analysis of § 5 down to equation (8) still holds when $M < 1$, so that, as before, if

$$z = e^g x, \quad \alpha = Ae^{i\gamma}, \quad q = 2\gamma + 1,$$

and

$$|\sin\gamma \log|z| - \cos\gamma \arg z| < \tfrac{1}{2}\pi A + \pi\cos\gamma - \pi|(q\cos\gamma - A)|,$$

then

$$\sum_{n=\gamma+1}^{\infty} \frac{\phi(n)e^{gn}}{\Gamma(\alpha n+1)} x^n = I_4 + \sum_{\mu=-p}^{p} F((xe^g)^{1/\alpha}\exp(2\pi i\mu/\alpha)),$$

where

$$I_4 = O(x^L) \text{ and } L < j + 1;$$

also

$$F(y) = \sum_{n=h+1}^{\infty} y^n \frac{\phi(n/\alpha)}{n!}.$$

Now we know that when $|s| \geqslant \gamma'$ and $|\arg(s/\alpha)| \leqslant \lambda + \frac{1}{2}\pi$, $\phi(s/\alpha)$ is analytic and

$$\phi(s/\alpha) = \exp(-\beta'\log s)\,.\left[a_0'' + \frac{a_1''}{s} + \frac{a_2''}{s^2} + \ldots + \frac{a_n''}{s^n} + R_n\right],$$

where

$$|a_n''| < A_1\rho^n\,.\,n!, \quad |R_n s^{n+1}| < A_2\sigma_1^n\,.\,n!.$$

Let $h$ be an integer greater than $\gamma'$ and let $\Theta$ be such that $h < -R(\Theta) < h + 1$; then, if $h_2$ denote any of the numbers 1, 2, ... $h$, we have

$$\frac{1}{s-h_2} = \frac{1}{s}\left[1 + \frac{h_2}{s} + \frac{h_2^2}{s^2} + \ldots + \frac{h_2^n}{s^n} + \frac{h_2^{n+1}}{s^{n+1}}\left\{1 + \frac{h_2}{s-h_2}\right\}\right],$$

so that, if $h < h_3 < -R(\Theta)$, then, when $|s| > h_3$

$$\frac{1}{s-h_2} = \frac{1}{s}\left[\bar{a}_0 + \frac{\bar{a}_1}{s} + \ldots + \frac{\bar{a}_n}{s^n} + \bar{R}_n\right],$$

where $|\bar{a}_n| < A_3\rho^n . n!$, $|\bar{R}_n s^{n+1}| < A_4\sigma_1^n . n!$, and $A_3$, $A_4$ are independent of $n$. (This follows, without difficulty, from Stirling's formula.)

Now if we have a finite number of asymptotic expansions with the same 'characteristics,' their product may be represented by an asymptotic expansion with the same characteristics*, that is to say that, when $|s| \geqslant h_3$, $|\arg s| \leqslant \lambda + \frac{1}{2}\pi$,

$$\frac{\phi(s/\alpha)}{s(s-1)\ldots(s-h)} = \exp\{(\beta' - h - 1)\log s\} . \left[a_0^{(1)} + \frac{a_1^{(1)}}{s} + \ldots + \frac{a_n^{(1)}}{s^n} + R_n^{(1)}\right],$$

where $|a_n^{(1)}| < A_1^{(1)}\rho^n . n!$, $|R_n^{(1)} s^{n+1}| < A_2^{(1)}\sigma_1^n . n!$.

Applying the expansion (2 $b$) to the function $\phi(s/\alpha) . \{s(s-1)\ldots(s-h)\}^{-1}$ instead of to the function† $\phi(s/\alpha)$, we get, when $R(s+\Theta) > 0$,

$$\frac{\phi(s/\alpha)}{s(s-1)\ldots(s-h)} = \frac{1}{(s+\Theta)^b}\left[b_0 + \frac{b_1}{M(s-\mu)+1} + \frac{b_2}{\{M(s-\mu)+1\}\{M(s-\mu)+2\}} + \ldots\right] \ldots(21),$$

where $|b_k| < H''\Gamma(k)\{\log(k+1)\}^V$, ($k > 1$, $H''$ independent of $k$),

and $R(b) < 0$, while $V$ is the integer such that $0 > R(b+V) \geqslant -1$, $\mu = -R(\Theta)$.

Using the notation of § 7, we put

$$h + 1 + \Theta = \theta, \quad R(h+1+\Theta) = c,$$

so that $0 < c < 1$.

It may be noted that the first $b - \beta' - h - 1$ of the coefficients $b_0, b_1, b_2, \ldots$ vanish; this is evident when we consider the effect of making $s \to \infty$ in (21).

From (9) and (21) it follows that

$$F(y) = y^{h+1}\sum_{n=0}^{\infty}\frac{y^n}{n!}\frac{1}{(n+\theta)^b}\left[b_0 + \frac{b_1}{M(n+c)+1} + \frac{b_2}{\{M(n+c)+1\}\{M(n+c)+2\}} + \ldots\right]$$

$$= y^{h+1}\sum_{m=0}^{k} b_m G_b(y;\theta, m; M, c) + y^{h+1}\sum_{n=0}^{\infty}\frac{B_{k,n}\,y^n}{(n+\theta)^b . n!},$$

where

$$|B_{k,n}| = \left|\left\{\frac{b_{k+1}}{\{M(n+c)+1\}\ldots\{M(n+c)+k+1\}} + \frac{b_{k+2}}{\{M(n+c)+1\}\ldots\{M(n+c)+k+2\}} + \ldots\right\}\right|$$

$$\leqslant \frac{|b_{k+1}|\,\Gamma(Mc+1)}{\Gamma(Mc+k+2)} + \frac{|b_{k+2}|\,\Gamma(Mc+1)}{\Gamma(Mc+k+3)} + \ldots.$$

*Qua* function of $k$ and $m$, $b_{k+m}$ is $O(\{\log(k+m+1)\}^V (k+m+1)^{-Mc-2})$, so that

$$B_{k,n} \text{ is } O\left(\sum_{m=1}^{\infty}\frac{\{\log(k+m+1)\}^V}{(k+m+1)^{Mc+1}}\right), \text{ i.e. } B_{k,n} \text{ is } O\left(\frac{\{\log(k+1)\}^V}{(k+1)^{Mc}}\right),$$

by Cauchy's condensation formula, since

$$Mc > 0 \text{ and } \{\log(k+m+1)\}^V (k+m+1)^{-Mc-1}$$

diminishes as $m$ increases when $k + m + 1 > \exp\{V(Mc+1)^{-1}\}$.

* See *Phil. Trans. Roy. Soc.*, vol. CCXI. A. (1911), pp. 279—313, § 4.

† This is legitimate since the former function satisfies the requisite conditions.

Since $$B_{k,n} \text{ is } O\left(\frac{\log(k+1)^{\Gamma}}{(k+1)^{Mc}}\right),$$

it follows that, for any fixed value of $y$ (i.e. any value not depending on $k$),

$$\sum_{p=0}^{\infty} \frac{B_{k,n} y^n}{(n+\theta)^p . n!} \to 0 \text{ as } k \to \infty ;$$

that is to say,

$$F(y) = y^{h+1} \sum_{m=0}^{\infty} b_m G_b(y;\, \theta, m;\, M, c).$$

Taking $\nu$ to be any fixed integer, we have the equation

$$F(y) = y^{h+1} \sum_{m=0}^{\nu} b_m G_b(y;\, \theta, m;\, M, c) + y^{h+1} \sum_{m=\nu+1}^{\infty} b_m G_b(y;\, \theta, m;\, M, c).$$

Now, by the theory of partial fractions, we may write

$$\frac{1}{\{M(n+c)+1\}\{M(n+c)+2\}\dots\{M(n+c)+m\}} = \sum_{s=1}^{m} \frac{p_{s,m}}{M(n+c)+s},$$

and also $$\frac{1}{M(n+c)+s} = \frac{1}{M}\left\{\frac{1}{n+1} + \frac{1-c-sM^{-1}}{(n+1)(n+2)} + \frac{(1-c-sM^{-1})(2-c-sM^{-1})}{(n+1)(n+2)(n+3)} + \dots\right\},$$

this series converging absolutely when $n+1 > 1-c-sM^{-1}$; this condition is certainly satisfied when $n \geqslant 0$.

That is to say, when $n \geqslant 0$, $0 < m \leqslant \nu$ ($\nu$ fixed), we may expand

$$\frac{1}{\{M(n+c)+1\}\{M(n+c)+2\}\dots\{M(n+c)+m\}}$$

into the absolutely convergent series

$$\frac{d_{1,m}}{n+1} + \frac{d_{2,m}}{(n+1)(n+2)} + \dots,$$

where $d_{k,m}$, *qua* function of $k$, is $O(\Gamma(k-c-M^{-1}))$; we may now justify the rearrangement of the series

$$\sum_{m=0}^{\nu} b_m G_b(y;\, \theta, m;\, M, c),$$

in the form $$\begin{aligned} b_0 G_b(y, \theta) &+ (b_1 d_{1,1} + b_2 d_{1,2} + \dots + b_\nu d_{1,\nu}) G_b(y, \theta, 1) \\ &+ (b_1 d_{2,1} + b_2 d_{2,2} + \dots + b_\nu d_{2,\nu}) G_b(y, \theta, 2) \\ &+ \dots, \end{aligned}$$

as in the somewhat similar work of § 5, so that we may write

$$\sum_{m=0}^{\nu} b_m G_b(y;\, \theta, m;\, M, c) = \sum_{m=0}^{\nu} d_m G_b(y;\theta, m) + \sum_{m=\nu+1}^{\infty} d'_m G_b(y;\, \theta, m),$$

where $d'_m$, *qua* function of $m$, is $O(\Gamma(m-c-M^{-1}))$.

Consequently

$$F(y) = y^{h+1} \sum_{m=0}^{\nu} d_m G_b(y;\, \theta, m) + \sum_{m=\nu+1}^{\infty} y^{h+1} \{d'_m G_b(y;\, \theta, m) + b_m G_b(y;\, \theta, m;\, M, c)\}.$$

[Since the coefficient of $y^{h+1+n}$ in the general term of the second summation, *qua* function of $n$, is

$$O\left(\frac{1}{(n+\theta)^p . (n+m)!}\right),$$

it is not difficult to see that the coefficients $d_m$ are those which occur* in the *formal* expansion:

$$\phi\left(\frac{n+h+1}{\alpha}\right)=\frac{(n+h+1)!}{n!\,(n+\theta)^b}\left\{d_0+\frac{d_1}{n+1}+\frac{d_2}{(n+1)(n+2)}+\dots\right\},$$

even though this expansion does not converge for any value of $n$.]

From the equation

$$F(y)=y^{h+1}\sum_{m=0}^{\nu} d_m G_b(y;\theta,m)+\sum_{m=\nu+1}^{\infty} y^{h-1}\{d'_m G_b(y;\theta,m)+b_m G_b(y;\theta,m;M,c)\}\ \dots(22),$$

we may deduce the asymptotic expansion of $F(y)$; for we have, by (6 $c$),

$$y^{h+1}\sum_{m=0}^{\nu} d_m G_b(y;\theta,m)=\frac{e^y}{y^{b-h-1}}\left[\sum_{m=0}^{\nu}\frac{T_m\Gamma(1-b)}{y^m}+o(|y|^{-\nu})\right]+O(|y|^h+|y^{1-\vartheta+h}|),$$

where
$$T_m=\sum_{n=0}^{m}\frac{d_m c_{n,m-n}}{\Gamma(1-b-n)},$$

and the coefficients $c_{n,m}$ are defined by the expansion

$$[\log(1-y)]^{b-1}(1-y)^{\vartheta-m-1}\equiv(-y)^{b-1}\sum_{n=0}^{\infty}c_{n,m}(-y)^m.$$

Also from (6 $a$) and (20) we have

$$\left|\sum_{m=\nu+1}^{\infty}\{d'_m G_b(y;\theta,m)+b_m G_b(y;\theta,m;M,c)\}\right|$$
$$<\sum_{m=\nu+1}^{\infty}|d'_m|\left\{\frac{K|e^y|}{|y^{b+b-1}|\,\Gamma(m+1)}+\frac{K|y^{1-b}|}{\Gamma(m+1)}+\frac{K\,.\,m}{\Gamma(m+1)\{\log m\}^{1-R(b)}}\right\}$$
$$+\sum_{m=\nu+1}^{\infty}|b_m|\left\{\frac{K\,U}{|y^{b-1}|\,\Gamma(Mc+m+1)}+\frac{K|y^{1-b}|}{\Gamma(Mc+m+1)}+\frac{K}{\Gamma(m+1)(\log m)^{1-R(b)}}\right\}.$$

Since $d'_m=O(\Gamma(m-c-M^{-1}))$, it follows that

$$\sum_{m=\nu+1}^{\infty}\frac{|d'_m|}{\Gamma(m+1)}=\sum_{m=\nu+1}^{\infty}O(m^{-1-c})<K \text{ since } c>0,$$

and
$$\sum_{m=\nu+1}^{\infty}\frac{|d'_m|\,m^c}{\Gamma(m+1)\{\log m\}^{1-R(b)}}=\sum_{m=\nu+1}^{\infty}\frac{1}{m\{\log m\}^{1-R(b)}}<K, \text{ since } R(b)<0.$$

Also since $b_m=O\{\Gamma(m)(\log m)^\Gamma\}$, it follows that

$$\sum_{m=\nu+1}^{\infty}\frac{|b_m|}{\Gamma(Mc+m+1)}=\sum_{m=\nu+1}^{\infty}O(m^{-1-Mc}\{\log m\}^\Gamma)<K, \text{ since } Mc>0,$$

and
$$\sum_{m=\nu+1}^{\infty}\frac{|b_m|}{\Gamma(m+1)\{\log m\}^{1-R(b)}}=\sum_{m=\nu+1}^{\infty}O(m^{-1}\{\log m\}^{\Gamma+R(b)-1})<K \text{ since } \Gamma+R(b)<0.$$

Consequently, if $l\leqslant\nu$,

$$\left|\sum_{m=\nu+1}^{\infty}\left\{d'_m G_b(y;\theta,m)+b_m G_b(y;\theta,m;M,c)\right\}\right|$$
$$<\frac{K|e^y|}{|y^{b-b-1}|}+K|y^{1-b}|+K+K\,U_y|y^{1-b}|\ \dots\dots\dots\dots(23),$$

where $U_y=1$ if $R(y)\leqslant 0$; if $R(y)\geqslant 0$, $U_y<(\sec\arg y)^l\,|y|^{-l}\,|e^y|+|e^{\frac{1}{2}y}|$ and $U_y<|e^y|$.

* The coefficients can be obtained successively by a limiting process.

That is to say, if $|\arg y| \leqslant \frac{1}{2}\pi - \Delta$, $\Delta > 0$,

$$\left|\sum_{m=\nu+1}^{\infty} \{d'_m G_b(y; \theta, m) + b_m G_b(y; \theta, m; M, c)\}\right| = O(|e^y y^{1-b-l}|) \quad \text{.........(23 a)},$$

while if $\pi \geqslant \arg y \geqslant \frac{1}{2}\pi - \Delta$,

$$\left|\sum_{m=\nu+1}^{\infty} \{d'_m G_b(y; \theta, m) + b_m G_b(y; \theta, m; M, c)\}\right| = O(|e^y y^L|) + O(y^L) \quad \text{......(23 b)},$$

where $L$ is determinate.

We deduce that, if $|\arg y| \leqslant \frac{1}{2}\pi - \Delta$, $F(y)$ possesses the asymptotic expansion

$$F(y) = \frac{e^y}{y^{b-h-1}}\left[\sum_{m=0}^{l-1} \frac{T_m \Gamma(1-b)}{y^m} + O(|y|^{-l+1})\right] \quad \text{.....................(24)},$$

while if $\pi \geqslant \arg y \geqslant \frac{1}{2}\pi - \Delta$,

$$F(y) = O(|e^y y^L|) + O(y^L) \quad \text{.................................(24 a)},$$

where $l$ is any fixed integer and $L$ is determinate.

12. From equation (8), quoted at the beginning of § 11, we can now see at once that if any one or more of the expressions $(xe^g)^{1/\alpha} \exp(2\pi i t/\alpha)$, where $t = -p, 1-p, \ldots p$, has its argument $\leqslant \frac{1}{2}\pi - \Delta$, the asymptotic expansion of $f(x)$ is given by the formula:

$$f(x) = \sum_{t=-p}^{p} \left[\frac{\exp(x^{1/\alpha} e^{g/\alpha + 2\pi i t/\alpha})}{(x^{1/\alpha} e^{g/\alpha + 2\pi i t/\alpha})^{b-h-1}} \times \left\{\sum_{n=0}^{\nu} \frac{T_n \Gamma(1-b)}{(x^{1/\alpha} e^{g/\alpha + 2\pi i t/\alpha})^n} + o(|x^{-n/\alpha}|)\right\}\right] \quad \text{.........(25)}.$$

The choice of the numbers $p$ and $q$ with the corresponding determination of the value of $\arg x$ is now made in precisely the same manner as in § 8, subsections (A)...(D).

Remembering that the first $1 + h + \beta' - b$ of the coefficients $b_0, b_1, \ldots$ vanish, we see that the first $1 + h + \beta' - b$ of the coefficients $T_0, T_1, \ldots$ vanish, and so it is convenient to take $\nu > 1 + h + \beta' - b$.

13. It is easy to see that the $n$th of the coefficients $T_0, T_1, \ldots$ (commencing with the first which does not vanish) is a linear combination of $a_0'', a_1'', \ldots, a''_{n-1}$, where

$$\phi(s/\alpha) = \exp(-\beta' \log s)\left[a_0'' + \frac{a_1''}{s} + \ldots\right];$$

and the only other arbitrary elements in $T_n$ are $\beta'$ and $\theta$; and it is evident that the expressions for $T_n$ will be *unaltered in form* if we supposed that the development

$$a_0'' + \frac{a_1''}{s} + \frac{a_2''}{s^2} + \ldots$$

were not asymptotic but convergent.

*That is to say, we may obtain the asymptotic expansion of $f(x)$ when $\phi(s)$ satisfies the conditions specified in § 2, by treating the development for $\phi(s/\alpha)$, viz.*

$$\exp(-\beta' \log s)\left\{a_0'' + \frac{a_1''}{s} + \frac{a_2''}{s^2} + \ldots\right\},$$

*as if it were not asymptotic but convergent*; and, in particular, by taking $\alpha = 1$, if the function $\chi(n + \theta)$ of § 1 be not analytic in the vicinity of $n = \infty$, but possess an asymptotic development, with grades equal to unity, valid for the range $|\arg n| < \frac{1}{2}\pi + \lambda$, $(\lambda > 0)$, Barnes' development* is still valid.

The asymptotic development which we have obtained seems to be an interesting example of the fact that in certain circumstances it is permissible to treat a subsidiary asymptotic development as if it were a convergent series.

* See § 1.

14. The theorem which we have proved may be enunciated formally as follows*:

Let $f(x)$ be an integral function defined by Taylor's series,

$$f(x) = c_0 + c_1 x + c_2 x^2 + \dots$$

and let it be possible to define a function $\phi(s)$ such that when $s$ is an integer

$$c_n = \frac{e^{gn}\phi(n)}{\Gamma(\alpha n+1)}, \qquad (|\arg\alpha| < \tfrac{1}{2}\pi,\ R(\alpha) \neq 0)$$

and $\phi(s)$ is such that when $|s| > \gamma'$, $|\arg s| \leqslant \frac{1}{2}\pi + \lambda$, $(\lambda > 0)$, $\phi(s/\alpha)$ is analytic†, and possesses the asymptotic expansion:

$$\phi(s/\alpha) = \exp(-\beta' \log s)\left[a_0'' + \frac{a_1''}{s} + \frac{a_2''}{s^2} + \dots + \frac{a_n''}{s^n} + R_n\right],$$

where $|a_n''| < A_1 \rho^n . n!$, $|R_n s^{n+1}| < A_2 \sigma_1^n . n!$ and $A_1$, $A_2$, $\rho$, $\sigma_1$ are independent of $n$.

Let $b$ be a number such that $R(b) < 0$ and $\beta' - b$ is zero or a positive integer; let $\Theta$ be such that $-R(\Theta) > \gamma'$ and let $h$ be the integer such that $h < -R(\Theta) < h+1$.

Let
$$\frac{\phi(s/\alpha)(s+\Theta)^b}{s(s-1)\dots(s-h)}$$

be expanded *formally* into the series of inverse factorials

$$\frac{e_0}{(s+1)(s+2)\dots(s+\beta'-b+h+1)} + \frac{e_1}{(s+1)(s+2)\dots(s+\beta'-b+h+2)} + \dots$$

(this development does not necessarily converge for any value of $s$).

Then, for the range of values of $x$ specified in § 8, $f(x)$ can be expanded into an aggregate of series which are asymptotic in the sense of Poincaré, thus,

$$f(x) = \sum_{t=-p}^{p}\left[\frac{\exp(x^{1/\alpha}e^{g/\alpha+2\pi i t/\alpha})}{(x^{1/\alpha}e^{g/\alpha+2\pi i t/\alpha})^{\beta'}} \times \left\{\sum_{n=0}^{\nu} \frac{U_n \Gamma(1-b)}{(x^{1/\alpha}e^{g/\alpha+2\pi i t/\alpha})^n} + o(|x|^{-\nu})\right\}\right],$$

where $\nu$ is any fixed integer and the coefficients $U_n$ are given by the series

$$U_n = \sum_{m=0}^{n} \frac{e_{n-m}}{\Gamma(1-b-m)} d_{m,n-m},$$

and the coefficients $d_{m,n}$ are given by the expansion

$$[\log(1-y)]^{b-1}(1-y)^{\Theta-n-\beta'-1} \equiv (-y)^{b-1}\sum_{m=0}^{\infty} d_{m,n}(-y)^m,$$

so that the coefficients $U_n$ can be calculated with sufficient labour, should they be required.

15. It is not difficult to shew that in the case of the generalised hypergeometric series

$${}_pF_q(a_1, a_2, \dots a_p;\ \rho_1, \rho_2, \dots \rho_q;\ x) \qquad (q+1>p),$$

the coefficients satisfy the conditions laid down for the functions discussed in this memoir. We thus obtain an independent proof of the results contained in Barnes' memoir on the generalised hypergeometric function‡.

Another class of functions which satisfy the conditions of this memoir is obtained by replacing the gamma functions involved in the generalised hypergeometric series by reciprocals of $G$-functions; it will be remembered§ that the $G$-function is an integral function satisfying the difference-equation $G(n+1) = \Gamma(n)G(n)$; and it is not difficult to construct other functions whose asymptotic expansions are given by the result stated in § 14.

* We have changed the notation slightly, in such a way that the first coefficient in the asymptotic expansion of $f(x)$ does not vanish.

† $c_n$ is supposed to be finite when $n$ is such that $\phi(s)$ is not analytic near $s = n$.

‡ *Proc. Lond. Math. Soc.*, ser. 2, vol. v. (1906), pp. 59—118.

§ See Barnes, *Quart. Journ. Math.* vol. xxxi. (1899), pp. 264—314.

# III. *The Hydrodynamical Theory of Lubrication with Special Reference to Air as a Lubricant.*

By W. J. HARRISON, M.A., Fellow of Clare College, Cambridge: Lecturer in Mathematics, Liverpool University.

[*Received* 11 May 1913. *Read* 19 May 1913.]

THE theory of the lubrication of surfaces moving relatively to one another and separated by a thin film of oil or other lubricant is one of considerable practical interest. The cognate problems are essentially hydrodynamical in their character, and have this interest, that they are among the few problems in the motion of viscous fluids which can be solved approximately for the case of large velocities. The theoretical work of Osborne Reynolds* and of Pétroff† is of extreme complexity, partly because Reynolds considered the case of an incomplete cylindrical bearing, and Pétroff introduced further complications into its form. It must be admitted that the forms of the bearing considered by these investigators are those which occur most frequently in practice, but their analysis is so complex and methods of approximation so laborious, that even the mathematician may fail to grasp the essential character of the results obtained, or to be expected. This fact alone is a valid reason for treating a simpler form of bearing. All results become simple, and the theory of lubrication can be elegantly illustrated by considering the case of a complete cylindrical bearing.

It was not till after I had completed my investigations in this case that I came across the very elaborate treatment by A. Sommerfeld‡ of the same problem. Our resulting formulae are identical. But the present treatment of the problem being somewhat different from Sommerfeld's, shorter and in one or two points more direct, will perhaps appeal more directly to experimenters.

A question is raised in the course of this paper as to the validity of experiments which have hitherto been made to determine the moment exerted by the traction of the lubricant on the journal. This point is of importance, as by means of this moment the nominal coefficient of friction of the journal is obtained.

In the latter part of this paper the method is extended to take account of lubrication by means of an elastic viscous fluid such as air. It was stated as long ago as 1885 by Hirn§, that under suitable circumstances air is the most perfect lubricant. In 1897 a series of very beautiful experiments was carried out by Prof. Kingsbury‖ on the lubrication of a cylindrical journal by air. The results he obtained, which are apparently accurate to a fair

* *Phil. Trans.*, 1886.

† *Mémoires de l'Acad. de St Pétersbourg*, VIII Sér. Classe Phys.-Math., Vol. X, 1900.

‡ *Zeits. für Mathematik*, Leipzig, 50, pp. 97—155, 1904.

§ *Engineering*, Jan. 30, 1885, p. 118.

‖ *Jour. Amer. Soc. Naval Engineers*, Vol. IX, 1897, p. 267.

degree, exhibit in certain details wide variation from those to be looked for in the case of lubrication by an incompressible liquid. I have not succeeded in obtaining an explicit solution of the differential equation determining the pressure in the film of air, but I have integrated it numerically by Runge's method, using the data of Kingsbury's experiments. The degree of approximation of theory to experiment is quite satisfactory. I have integrated the differential equation in the case of plane surfaces, and give some results below which exhibit more clearly the very marked effects of the compressibility of the air on the magnitude and distribution of the pressure.

But apart from the new results obtained, this paper will serve the useful purpose of recalling attention to Sommerfeld's work. A subsequent paper by A. G. B. Michell* is also worth attention and will be referred to below.

It might be in place to remark here that I have obtained some results and have work in hand treating of cases in which the influence of variable speed and variable load on the lubrication of a cylindrical bearing is taken into account.

### *Case of Incompressible Liquid.*

In proceeding to determine the equations which give the motion and the pressure of a film of liquid separating two surfaces moving relatively to one another, it is to be observed that the inertia terms can be neglected as well as the effect of gravity, since forces depending on these terms are negligible compared with the internal stresses arising from the rapid shearing of the liquid. Again, on account of the thinness of the film its curvature can be neglected, and therefore the same equations hold whether the surfaces are plane or cylindrical. Sommerfeld has transformed the equation $\nabla^4\psi = 0$, which is satisfied by the stream function $\psi$, from Cartesian coordinates $(x, y)$ to polar coordinates $(r, \theta)$. He proceeds to use essentially the same method of approximation as employed by Osborne Reynolds. The only result of this transformation is to introduce relatively unimportant terms, as will be seen below.

The coordinate $x$ will be measured along the moving surface in the direction of motion, the coordinate $y$ normal to this surface. The motion is steady and will be assumed to be two-dimensional.

If $u$, $v$ be the component velocities at any point in the liquid, $p$ the pressure, the equations of motion are

$$\frac{\partial p}{\partial x} = \mu\nabla^2 u \qquad (1),$$

$$\frac{\partial p}{\partial y} = \mu\nabla^2 v \qquad (2),$$

where $\mu$ is the coefficient of viscosity, and the equation of continuity is

$$\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} = 0 \qquad (3).$$

* *Zeitschrift für Mathematik*, Leipzig, 52, 1905, pp. 123—137.

The boundary conditions are

$$\left.\begin{array}{llll} u=U, & v=0, & \text{when} & y=0 \\ u=0, & v=0, & \text{when} & y=h \end{array}\right\} \quad\ldots\ldots(4),$$

where $h$ is the variable distance between the surfaces and is a function of $x$.

Since the surfaces are nearly parallel $v$ will be small compared with $u$, and the rate of variation of $u$ in the direction of $x$ will be very small compared with its rate of variation in the direction of $y$.

Accordingly equations (1) and (2) become

$$\frac{\partial p}{\partial x}=\mu\frac{\partial^2 u}{\partial y^2} \quad\ldots\ldots(5),$$

$$\frac{\partial p}{\partial y}=0 \quad\ldots\ldots(6).$$

From (6) it is seen that $p$ is independent of $y$, and (5) is then integrable, giving

$$u=\frac{1}{2\mu}\frac{\partial p}{\partial x}y(y-h)+U\frac{h-y}{h},$$

in which the constants of integration have been adjusted to suit the boundary conditions (4).

Now from (3) $$\int_0^h \frac{\partial u}{\partial x}dy=-\left[v\right]_0^h=0.$$

Hence $$\frac{\partial}{\partial x}\left(h^3\frac{\partial p}{\partial x}\right)=6\mu U\frac{\partial h}{\partial x},$$

which gives $$h^3\frac{\partial p}{\partial x}=6\mu U(h-h_1),$$

where $h_1$ is the value of $h$ at a place of stationary pressure.

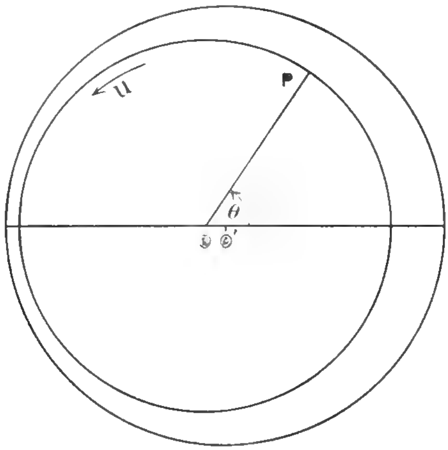



Fig. 1.

In the case of a cylindrical journal, let $O$ be its centre, $a$ its radius; $O'$ the centre of the outer bearing, $a+\eta$ its radius. Let $OO'=c\eta$, $0<c<1$. Then the radial distance between the two surfaces can be put into the form $h = \eta(1 + c\cos\theta)$, where $\theta$ is the angle $POO'$. Hence, writing $x = a\theta$, $h_1 = \eta(1 + c\cos\theta_1)$, we obtain

$$\frac{dp}{d\theta} = \frac{6\mu Uac(\cos\theta - \cos\theta_1)}{\eta^2(1 + c\cos\theta)^3}.$$

This equation can be integrated out in finite terms, but Osborne Reynolds, owing to the fact that he was solving the case of an incomplete cylindrical bearing, found it more convenient to expand in a series of ascending powers of $c$, and integrate term by term.

It is easily found that

$$p = C - \frac{6\mu Ua}{(1-c^2)^2\eta^2}\left[\frac{c\sin\theta}{(1+c\cos\theta)^2}\{-(1-c^2)(1+c\cos\theta) + \tfrac{1}{2}(1+c\cos\theta_1)(1-c^2-3[1+c\cos\theta])\}\right.$$
$$\left. + (1-c^2)^{-\frac{1}{2}}\{2(1-c^2) - (1+c\cos\theta_1)(2+c^2)\}\tan^{-1}\left\{\sqrt{\frac{1-c}{1+c}}\tan\frac{\theta}{2}\right\}\right].$$

Now $p$ must be a single-valued function of $\theta$, hence

$$3c + (2 + c^2)\cos\theta_1 = 0 \quad \text{..........................................(7)}.$$

This equation determines the positions of the max.-min. values of the pressure. The remaining constant can be determined if the pressure is known at any one point. It is to be noticed that equation (7) does not restrict the value of $c$ except to the range $-1 < c < 1$.

Substituting the value of $\cos\theta_1$ so obtained, we have

$$p = C + \frac{6\mu Uac\sin\theta(2 + c\cos\theta)}{\eta^2(2 + c^2)(1 + c\cos\theta)^2}.$$

The positions of maximum and minimum pressure are equidistant from the point of nearest approach, and the one value rises above the value of the pressure at that point by as much as the other value falls below it.

It remains to calculate the resultant forces and couples acting on the surfaces of the journal and bearing due to the pressure and traction of the liquid.

The component of force exerted on the journal due to the pressure $p$ is $R$ acting downwards through $O$ (see figure 1). The component along $OO'$ vanishes, and

$$R = \int_0^{2\pi} p\sin\theta\, a\, d\theta = \frac{12\pi\mu Ua^2c}{\eta^2(2 + c^2)(1 - c^2)^{\frac{1}{2}}}.$$

The component of force exerted on the journal by the traction $f$ (measured in a direction opposite to that of the rotation) is $S$ acting along $OO'$.

Now
$$f = -\mu\left(\frac{\partial u}{\partial y}\right)_{y=0}$$
$$= \left[\frac{\mu U}{h} + \frac{1}{2}h\frac{\partial p}{\partial x}\right]_{y=0}$$
$$= \frac{\mu U}{\eta}\left[\frac{4}{1 + c\cos\theta} - \frac{6(1 - c^2)}{(2 + c^2)(1 + c\cos\theta)^2}\right].$$

Hence
$$S = \int_0^{2\pi} f\sin\theta\, a\, d\theta = \frac{4\pi\mu Ua}{\eta c}\left[2 - \frac{4 - c^2}{(2 + c^2)(1 - c^2)^{\frac{1}{2}}}\right].$$

The couple exerted on the journal is

$$M = \int_0^{2\pi} fa^2 d\theta = \frac{4\pi\mu U a^2 (1 + 2c^2)}{\eta (2 + c^2)(1 - c^2)^{\frac{1}{2}}},$$

being measured in a direction opposite to that of the rotation.

Now if $f'$ be the tangential traction on the surface of the bearing, measured in the direction of rotation,

$$f' = -\mu \left(\frac{\partial u}{\partial y}\right)_{y=h}$$

$$= \mu \left[\frac{U}{h} - \tfrac{1}{2} h \frac{\partial p}{\partial x}\right].$$

Hence on the outer surface the corresponding forces and couples are

$R'$ (acting upwards through $O'$) $= R$,

$$S' = \frac{4\pi\mu U a}{\eta c}\left[1 - \frac{2(1 - c^2)^{\frac{1}{2}}}{2 + c^2}\right],$$

$$M' = \frac{4\pi\mu U a^2 (1 - c^2)^{\frac{1}{2}}}{\eta (2 + c^2)}.$$

In the first place it is to be noticed that $S$ and $S'$ are not equal and opposite, but these components are of a smaller order than $R$ and $R'$, and will therefore be neglected; a closer approximation would establish their equality—in fact, in Sommerfeld's work they are shown to vanish. The expressions for $R$ and $M$ given above are the same as those obtained by Sommerfeld.

But the inequality of $M$ and $M'$ is essential. Taking into account the fact that $R$ acts at the point $O$, and $R'$ at the point $O'$, it is easily verified that the force system $(R, M)$ is balanced by the system $(R', M')$.

The ratio of $M$ to $M'$ for different values of $c$ is as follows:

| | | | | | |
|---|---|---|---|---|---|
| $c =$ | 0 | ·1 | ·4 | ·6 | ·9 |
| $M/M' =$ | 1 | 1·03 | 1·51 | 2·69 | 13·8 |

Now this difference between $M$ and $M'$ is of very considerable importance, and so far as I am aware has not been previously noticed.

The following conventional terms are usually adopted:

$$\text{Nominal load} = \frac{\text{load per unit length of bearing}}{\text{diameter}},$$

$$\text{Nominal friction} = \frac{\text{couple due to traction}}{\text{radius} \times \text{diameter}},$$

$$\text{Coefficient of friction} = \frac{\text{nominal friction}}{\text{nominal load}}.$$

Hence the coefficient of friction $\lambda$ for the journal is given by

$$\lambda = \frac{M}{R \cdot a} = \frac{\eta (1 + 2c^2)}{3ac};$$

for the bearing

$$\lambda' = \frac{M'}{R' a} = \frac{\eta (1 - c^2)}{3ac}.$$

$\lambda$ has a minimum value $\dfrac{2\sqrt{2}\eta}{3a}$ when $c = 1/\sqrt{2}$; $\lambda'$ has no minimum value except zero. When $\lambda$ has its minimum value, $\lambda = 4\lambda'$.

It is immediately obvious that, if the coefficient of friction be determined by experiment, it is important to know what is the couple which is measured. It may be safely assumed that in experiments hitherto conducted the object has been to measure the couple exerted on the bearing, the direct measurement of the couple exerted on the journal being out of the question. But if this is so the actual coefficient of friction for a given journal is greater than that derived from experiments made upon it, unless the speed be sufficiently great to make $c$ comparatively small.

It is important to observe that there are in reality three force systems under consideration which are equivalent but not identical. The force $R$ and couple $M$ exerted on the journal by the liquid, the force $R'$ and couple $M'$ exerted by the liquid on the bearing, the force $R''$ and couple $M''$ applied to the bearing to keep it in position. Now $R = R' = R''$, in magnitude, but $M$ and $M'$ are always unequal since the lines of action of $R$ and $R'$ are not the same, and the magnitude of $M''$ depends on the line of action of $R''$. It is necessary that $M''$ should be equal to $M$, if the correct coefficient of friction is to be obtained by experiment, and accordingly the line of action of $R''$ must be adjusted so as to pass through the centre $O$ of the journal, which is a variable point depending on the load and velocity. Since the relative magnitudes of $Ra$ and $M$ are so disproportionate, the slightest error in this adjustment, even if the line of action of $R''$ be only $\frac{1}{1000}$ of an inch out, causes a considerable percentage error in the measured moment. In the case of incomplete cylindrical bearings, for which the arc of contact may be 120° or less, $M$ and $M'$ will be more nearly equal. But there is the same possibility that $M''$ will be equal to neither the one nor the other.

Suppose now the line of action of $R''$ to act at a constant distance $x$ from $O'$ towards $O$, i.e. we suppose the line of action of the applied force to be independent of the velocity of the journal and of the load applied to it.

Then $$M' = M'' - R''x,$$

hence $$\lambda'' = \lambda' + x/a,$$

where $\lambda''$ is the apparent observed coefficient of friction ($x/a$ may be very small and yet comparable with $\lambda'$, as seen previously for $x = c\eta = OO'$).

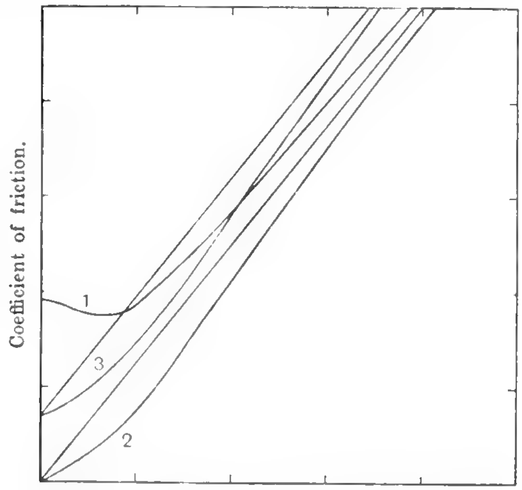



Fig. 2.

The variations of $\lambda$, $\lambda'$, $\lambda''$ with the speed are exhibited in figure 2. The graphs are not drawn quite accurately to scale, but they represent closely the behaviour of the three coefficients. Curve 1 shows the variation of $\lambda$ with its minimum value at $c = 1/\sqrt{2}$; curve 2 exhibits $\lambda'$, and this curve displaced a constant distance along the ordinates into the position of curve 3 exhibits $\lambda''$. The displacement of curve 2 to produce curve 3 may be either upwards or downwards.

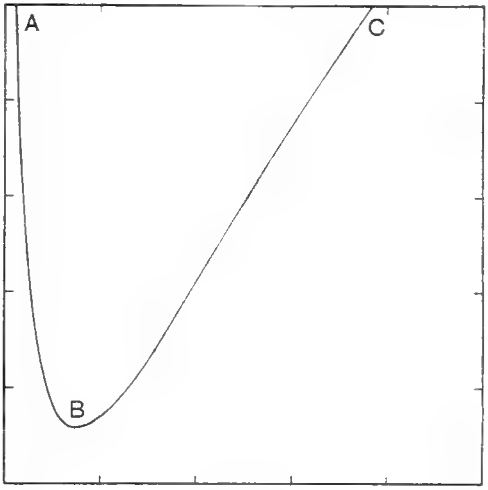



Fig. 3.

In figure 3 a typical curve is drawn showing the variation of the coefficient of friction, as observed, with the speed. It would appear that the part of the curve $AB$ represents the effect of imperfect lubrication owing to insufficient speed. The remainder of the curve $BC$ might very well be a part of any one of the curves shown in figure 2, or roughly approximate to one of them. But there is a good reason why it cannot be associated in general with curve 1, which exhibits the true coefficient of friction. For it is a fact worthy of attention that the observed minimum value of the coefficient of friction may be less than the minimum calculated value of $\lambda$; this is so in the case of Kingsbury's data. It is, moreover, clear that he measured the value of $M'$, and not that of $M$. He measured the couple exerted on the journal, while it was kept at rest and the outer bearing rotated, by the torsion of a wire fastened to a point on its axis.

It may be considered as moderately certain that experiments hitherto conducted cannot be relied on to give more than a rough estimate of the true coefficient of friction at fairly low speeds. It is hoped that some attention may be given to the points here raised in future experiments. Sufficient is now known about the theory of lubrication for all practical purposes, so that such refinements as are here proposed have for their object merely the possibility of obtaining data on which to base a closer comparison between theory and experiment.

If it is affirmed that the comparison which is presented in Osborne Reynolds' paper is sufficiently close, two objections to that comparison may be raised. ($a$) Pétroff has pointed out what seems to be a serious mistake in sign, which is carried on in subsequent operations. ($b$) Osborne Reynolds was compelled first to estimate the distance between the bearing surfaces by means of his theory and then make a final comparison between theory and experiment.

Such a course as that indicated in (*b*) is practically inevitable except in the case of a closed cylindrical bearing; in other cases the difficulty of obtaining by observation the relative positions of journal and bearing is great. Now this relative position of the two surfaces has been determined by Kingsbury in the case of an air-lubricated journal. Thus his data allow of a complete comparison. To a consideration of these experiments we now proceed.

It may be added that Sommerfeld gives a very complete treatment of the lubrication of a cylindrical journal by an incompressible liquid, which is further illuminated by a number of curves showing the relations between the various quantities $M$, $R$, $\lambda$, $c$, $U$ which enter into the theory. One such curve is reproduced in curve 1 of figure 2.

## *Case of Elastic Fluid.*

Some of the most accurate experiments which have been made on the lubrication of a complete cylindrical journal were those conducted by Prof. Kingsbury, using air as a lubricant.

He investigated the distribution of pressure round the journal and along its length, the point of closest approach and the magnitude of the shortest distance, and also the moment of the friction exerted on the bearing. But, regarded generally, the results show a marked divergence from the state of affairs indicated by the theory for an incompressible lubricant. This divergence is of course due to the elasticity of the air, and presents a somewhat interesting problem for investigation. The main points of this difference will be found exhibited in the table below, in which the air is treated as inelastic, having the density of atmospheric air, and average pressure equal to atmospheric pressure $\Pi$. All quantities are expressed in foot, lb., second units. The diameter of the journal was $\frac{1}{2}$ ft., and the difference of the radii $\frac{2}{3} \times 10^{-4}$ ft. $U$ is the velocity at the surface of the journal. The pressure $p$ is of course given in poundals per square foot. The first row of the double sets of data comprises those obtained by observation, the second row those calculated from the theory already developed in this paper.

| | | | |
|---|---|---|---|
| Revs. per min. | 230 | 805 | 1730 |
| $U$ | 6·02 | 21·07 | 45·29 |
| $c$ | ·39 | ·1625 | ·091 |
| $(p_{max}-\Pi)10^{-4}$ | 1·12 | ·63 | ·63 |
| | 1·16 | 1·44 | 1·69 |
| $(p_{min}-\Pi)10^{-4}$ | − ·85 | − ·68 | − ·73 |
| | −1·16 | −1·44 | −1·69 |
| $(p_{max}-p_{min})10^{-4}$ | 1·97 | 1·31 | 1·36 |
| | 2·32 | 2·88 | 3·38 |
| $(\theta_{max}-180°)$ | −28′ | −43′ | −43′ |
| | −57′ | −76′ | −82′ |
| $(\theta_{min}-180°)$ | 51′ | 96′ | 129′ |
| | 57′ | 76′ | 82′ |
| $\theta_{min}-\theta_{max}$ | 79′ | 139′ | 172′ |
| | 114′ | 152′ | 164′ |

It is clearly seen that the facts which need explanation are these: (1) the range of the pressure decreases and seems to approach more or less to a definite limit as the velocity increases; (2) the position of the maximum pressure is displaced nearer to the point of closest approach; (3) the position of the minimum pressure is displaced further from that point.

*Steady Flow of Air under Pressure.*

In the same way that the steady flow of liquid under pressure between two parallel planes is introductory to the problems of lubrication, so the flow of air in the same case serves as an introduction to our extended theory*.

Consider the flow of air in two dimensions between the parallel planes $y = \pm d$. It will be assumed as an approximation that the equations of motion and continuity are

$$\frac{\partial p}{\partial x} = \mu \frac{\partial^2 u}{\partial y^2} \quad \text{.........(8)},$$

$$\frac{\partial p}{\partial y} = 0 \quad \text{.........(9)},$$

$$\frac{\partial (\rho u)}{\partial x} = 0 \quad \text{.........(10)}.$$

The conditions are such that the relation between the pressure and the density is

$$p = k\rho.$$

Hence from (10) $\rho u$ or $pu$ is a function of $y$ only. Also from (9) $p$ is independent of $y$, and therefore from (8)

$$u = -\frac{1}{2\mu}\frac{dp}{dx}(d^2 - y^2) \quad \text{.........(11)},$$

taking into account the condition $u = 0$ at $y = \pm d$.

It follows that $p\dfrac{dp}{dx}$ is a function of $y$ only; accordingly put $p^2 = ax + b$.

Let $p_0$ be the pressure at $x = 0$, $p_1$ the pressure at $x = l$. Then

$$p^2 = p_0^2 - (p_0^2 - p_1^2)\, x/l.$$

Thus the velocity at any point is given by

$$u = \frac{(p_0^2 - p_1^2)(d^2 - y^2)}{4\mu l \{p_0^2 - (p_0^2 - p_1^2)\, x/l\}^{\frac{1}{2}}}.$$

For a circular tube of radius $a$ and length $l$, it is easily shown that†

$$u = \frac{(p_0^2 - p_1^2)(a^2 - r^2)}{8\mu l \{p_0^2 - (p_0^2 - p_1^2)\, z/l\}^{\frac{1}{2}}},$$

$r$ being the distance from the axis, $z$ the distance along the tube.

* In addition it has some bearing on the researches of Prof. A. H. Gibson on the flow of air through pipes. The formula obtained below has also been arrived at by G. F. C. Searle by a different method, *Proc. Camb. Phil. Soc.*, Vol. XVII, Pt II, 1913.

† Cf. Lamb, *Hydrodynamics*³, §§ 338, 339.

The mass of air crossing any section of the tube per unit time is

$$\int_0^a \rho u \, 2\pi r \, dr = \int_0^a (a^2 - r^2) \, 2\pi r \, dr \Big/ \left\{ \frac{8\mu k l}{p_0^2 - p_1^2} \right\}$$

$$= \frac{\pi a^4 (p_0^2 - p_1^2)}{16 \mu k l}$$

$$= \frac{\pi a^4 (p_0 - p_1)}{8\mu l} \frac{\rho_0 + \rho_1}{2}.$$

This is the formula assumed by Prof. Gibson *.

It remains to investigate the order of the terms neglected in the equations (8), (9), (10). Using the data of Gibson's experiments for velocities below the critical value, I find the following approximations, taking $\frac{dp}{dx}$ as unity:

$$\mu \frac{\partial^2 u}{\partial x^2} = 10^{-12}, \quad \mu \frac{\partial^2 u}{\partial x \partial y} = 10^{-6}, \quad \rho u \frac{\partial u}{\partial x} = 10^{-3}.$$

It would appear therefore from the order of the quantities neglected that in all ordinary cases of stream-line flow in pipes the approximation assumed is quite accurate. In fact, long before the velocities are such as to cause the approximation to break down turbulent motion supervenes.

In the same way the use of these approximations in the theory of lubrication could be justified. But it must be pointed out that, as a matter of fact, the approximations necessarily break down immediately in the neighbourhood of stationary pressure. But the effect of this failure is probably negligible, since it is confined to a very small part of the lubricant.

### *Motion of a Film of Air between Two Surfaces, one of which is in Motion.*

It is necessary to solve equations (8), (9), (10), subject to boundary conditions (4).

Equation (11), above, becomes

$$u = \frac{1}{2\mu} \frac{dp}{dx} y (y - h) - \frac{U}{h} (y - h).$$

Hence

$$\Big[ p v \Big]_{y=0}^{y=h} = - \int \frac{\partial}{\partial x} (p u) \, dy$$

$$= \frac{\partial}{\partial x} \left[ p \left( - \frac{h^3}{12\mu} \frac{dp}{dx} + \tfrac{1}{2} U h \right) \right].$$

But $v = 0$ at $y = 0$, $y = h$, and accordingly

$$- \frac{h^3}{\mu} \frac{dp}{dx} + 6 U h = \frac{k}{p} \quad \text{..........................................(12)},$$

where $k$ is a constant. It may be as well to remark that for a gas $\mu$ is independent of the pressure.

* *Proc. Roy. Soc.*, Vol. LXXX, p. 114, 1908.

*Lubrication of a Cylindrical Journal by Air.*

Using the same notation as in the previous part of the paper, equation (12) becomes

$$\frac{dp}{d\theta} = \frac{6\mu Ua}{\eta^2(1+c\cos\theta)^2}\left[1 - \frac{k}{\eta(1+c\cos\theta)p}\right] \quad \text{.......................(13).}$$

This equation I have not been able to integrate in any form convenient for calculation. But equation (12) is easily integrable in the case of inclined plane surfaces. I have solved a number of cases of motion between such surfaces to illustrate the effect of elasticity. These results have an interest of their own, and I shall present them later. But since making these calculations, which were intended to explain roughly the discrepancies between theory and experiment in the case of Kingsbury's data, I have integrated equation (13) numerically by Runge's method of numerical integration, using the data of Kingsbury's experiments. These numerical solutions will be found shown by curves 2 in figures 4, 5, 6. The ordinate represents pressure ($6{\cdot}83 \times 10^4$ is atmospheric pressure), the abscissa is $\theta$, the angle $POO'$ in figure 1. Curves 3 give the distribution of pressure round the journal observed by Kingsbury; curves 1 give the pressure on the supposition that air is inelastic. It is seen that the extended theory goes far towards the explanation of the discrepancies referred to above.

It will be noticed that in figures 4 and 6 the observed and calculated maxima and minima of the pressure are in good agreement. As regards the maximum pressure the agreement is not so good in figure 5. In connection with the differences which still remain it needs to be pointed out that the theory is based on the assumption of an infinitely long bearing. In particular the differences in position of max.-min. pressures in curves 2 and 3 are to be attributed partly to the finite length of the bearing. Michell, whose paper has been referred to above, has investigated the influence of finite length in the case of inclined plane surfaces. He states that he was unable to solve the same problem for a cylindrical journal.

It might be as well to add a few words on the numerical solution of equation (13). It was found necessary to divide the range of 360° up into intervals of 10° or 20°, according to

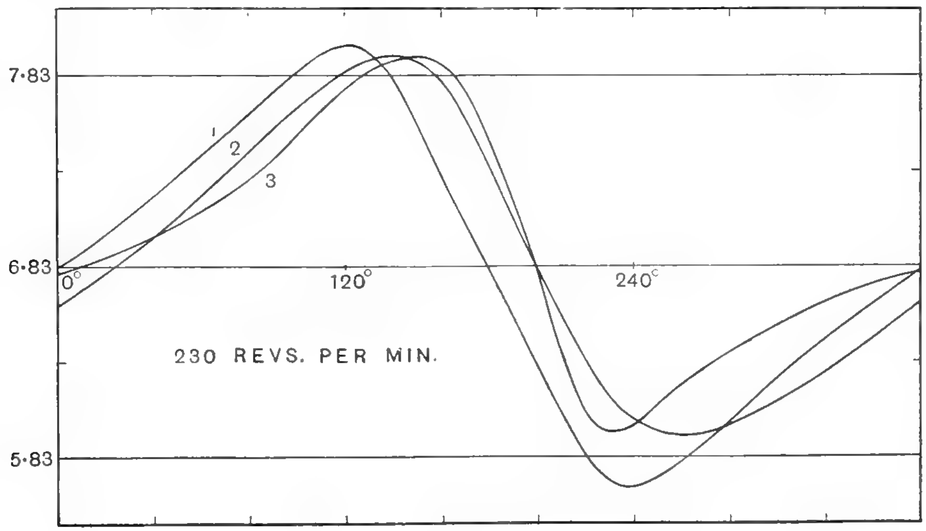



Fig. 4.

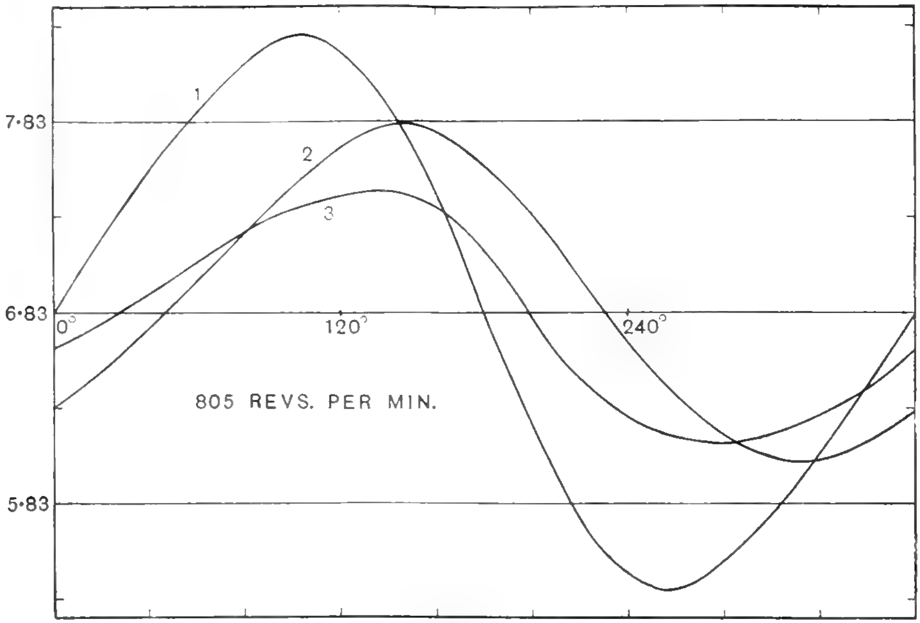



Fig. 5.

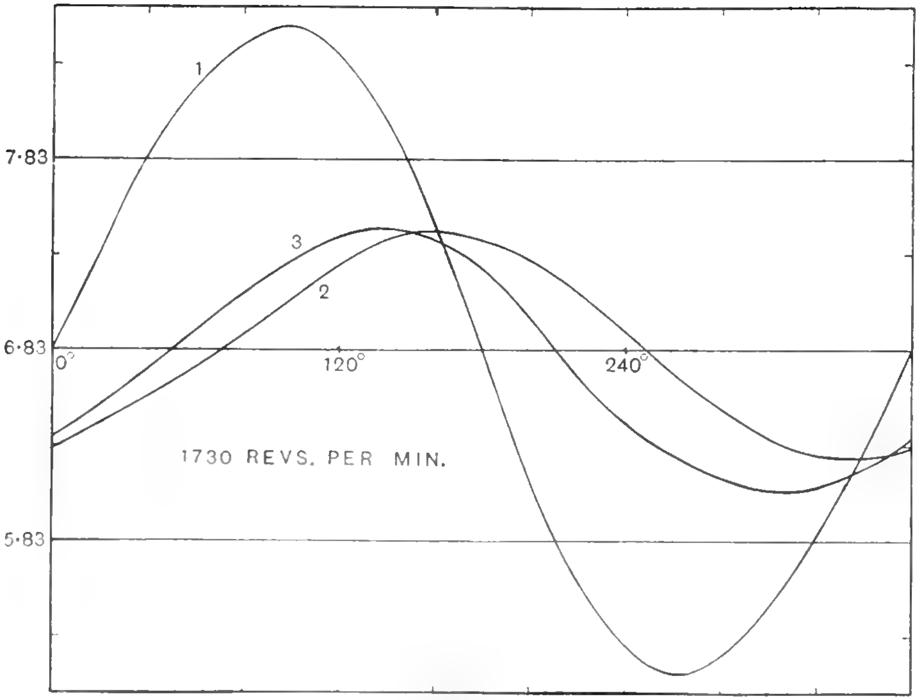



Fig. 6.

the position in the range. Not only did the number of intervals make the calculation very laborious (a single complete solution involved the writing down of 12,000 digits), but it was further necessary to provisionally assume the values of two constants before an evaluation could be commenced. These two constants were practically the pressure and its gradient at the beginning of the first interval. The tests as to the correctness of these assumed values are (1) the pressure at $\theta = 0^\circ$ must be equal to that at $\theta = 360^\circ$, (2) the total mass of air must be constant, or

$$\int_0^{2\pi} (p - \Pi)(1 + \cos\theta)\, d\theta = 0.$$

The assumed values of the constants have to be varied until these conditions are satisfied. To give further details here would take up too much space, but it is believed that the curves obtained are quite accurate.

*Case of Inclined Planes.*

Consider two plane surfaces of which the upper is fixed and is of unit breadth, inclined so that their greatest distance apart is $d_0$, and their least distance apart $d_1$. In the first place let the motion of the lower plane be as indicated in figure 7. This case has been treated by Osborne Reynolds for incompressible liquid, and he virtually obtained the following results:

Position of maximum pressure $\quad x_1 = d_0/(d_0 + d_1)$,

Distance between planes at this point $\quad h_1 = 2d_0d_1/(d_0 + d_1)$,

Maximum pressure given by $\quad p_{\text{max.}} - \Pi = \dfrac{3\mu U (d_0 - d_1)}{2d_0d_1(d_0 + d_1)}$.

In the extension to the case of elastic fluid we write $h = d_0 - bx$, where $b = d_0 - d_1$, in equation (12), which can then be written

$$\frac{bh^3}{\mu}\frac{dp}{dh} + 6Uh = \frac{k}{p}.$$

Put $ph = w$, and we have

$$\frac{\frac{b}{\mu} w\, dw}{\frac{b}{\mu} w^2 - 6Uw + k} - \frac{dh}{h} = 0.$$

The form of the integral depends on the sign of $\mu kb - 9U^2\mu^2 = K^2$, and for our present purpose this will be found to be positive within the required range of $U$.

Hence $$\tfrac{1}{2}\log(bw^2/\mu - 6Uw + k) + \frac{3\mu U}{K}\tan^{-1}\frac{wb - 3\mu U}{K} - \log h = C \quad \text{...........(14).}$$

Now $p = \Pi$ for $x = 0$, $h = d_0$; $x = 1$, $h = d_1$.

Hence substituting and subtracting the two equations so obtained, we have an equation from which to determine $k$; $C$ can then be determined. To find the value and position of the maximum pressure, we notice that when $\dfrac{dp}{dx} = 0$,

$$6\mu Uph = k \quad \text{.........................................(15).}$$

Hence the position $x_1$ of the maximum pressure is given by the equation

$$\log\frac{d_0 - bx_1}{d_0} = \tfrac{1}{2}\log\frac{2(k/6U)^2}{b\Pi^2 d_0^2/\mu - 6U\Pi d_0 + k} + \frac{3\mu U}{K}\left[\tan^{-1}\frac{bk/6U - 3\mu U}{K} - \tan^{-1}\frac{b\Pi d_0 - 3\mu U}{K}\right] \quad \text{...........(16).}$$

Having found $x_1$ from this equation the value of the maximum pressure follows at once from equation (14).

The following tables exhibit the results of calculations, for the purpose of which it was taken that $\Pi = 7 \times 10^4$ (15 lbs. per sq. in. approximately), $\mu = 10^{-5}$.

The second row of figures in each case comprises those obtained from Osborne Reynolds' formulae.

$d_0 = 10^{-4}$, $d_1 = \cdot 4 \times 10^{-4}$

| $U$ | 5 | 10 | 20 |
|---|---|---|---|
| $(p_{max.} - \Pi)\, 10^{-4}$ | ·80<br>·80 | 1·58<br>1·61 | 2·99<br>3·22 |
| $x_{max.}$ | ·75<br>·71 | ·77<br>·71 | ·81<br>·71 |

$d_0 = \cdot 8 \times 10^{-4}$, $d_1 = \cdot 6 \times 10^{-4}$

| $U$ | 5 | 10 | 20 |
|---|---|---|---|
| $(p_{max.} - \Pi)\, 10^{-4}$ | ·22<br>·22 | ·42<br>·47 | ·80<br>·89 |
| $x_{max.}$ | ·61<br>·57 | ·63<br>·57 | ·69<br>·57 |

Thus the effect of the elasticity is to force the position of maximum pressure nearer to the narrower end, and to decrease its value.

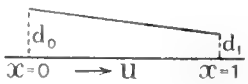


Fig. 7.

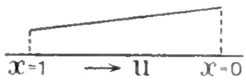


Fig. 8.

We proceed to consider now the case of motion indicated by figure 8. Similar equations to (14) and (16) are obtained, except that the $\tan^{-1}$ is replaced by a logarithmic form. The following are the results obtained by calculation.

$d_0 = 10^{-4}$, $d_1 = \cdot 4 \times 10^{-4}$

| $U$ | 5 | 10 | 20 |
|---|---|---|---|
| $(p_{min} - \Pi)\, 10^{-4}$ | −·77<br>−·80 | −1·44<br>−1·61 | 2·33<br>−3·22 |
| $x_{min}$ | ·67<br>·71 | ·63<br>·71 | ·53<br>·71 |

$d_0 = \cdot 8 \times 10^{-4}$, $d_1 = \cdot 6 \times 10^{-4}$

| $U$ | 5 | 10 | 20 |
|---|---|---|---|
| $(p_{\text{min.}} - \Pi)\,10^{-4}$ | $-\cdot 22$<br>$-\cdot 22$ | $-\cdot 37$<br>$-\cdot 47$ | $-\cdot 75$<br>$-\cdot 89$ |
| $x_{\text{min.}}$ | $\cdot 53$<br>$\cdot 57$ | $\cdot 52$<br>$\cdot 57$ | $\cdot 42$<br>$\cdot 57$ |

The effect of elasticity is to displace the position of minimum pressure in the direction of motion and to decrease the total range of pressure. It is interesting to notice that, in accordance with expectation, although for an incompressible liquid with the reversed velocity the fall of the minimum pressure below atmospheric pressure is the same as the rise of the maximum pressure above, yet the effect of elasticity is to make the rise of the maximum pressure above atmospheric pressure greater than the fall of the minimum pressure below it when the velocity is reversed.

As these cases of the lubrication of inclined plane surfaces were solved with a view to illustrating the differences between Kingsbury's data and calculations based on the theory for an incompressible liquid, it was necessary to obtain some solution in which both a maximum and a minimum pressure should appear. The case indicated in figure 9 suggested itself as instituting a fairly close comparison, the dimensions being chosen so as to agree as closely as possible with the dimensions of the cylindrical bearing. In fact the solutions of this case agree very well with the observed facts, and although this comparison is no longer necessary, the following solutions throw further light on the behaviour of an elastic lubricant.

Consider now the case of two inclined plane surfaces placed as in figure 9, with respect to a third moving plane surface. For an incompressible liquid the pressures at $A$, $B$, $C$ are all

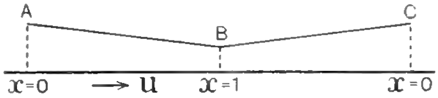



Fig. 9.

equal and the max.-min. pressures and their positions are the same as in the two separate cases, which have been already discussed. But this is not so if the fluid be elastic. In this case, for instance, the pressure at $B$ may exceed the maximum pressure in $AB$ in the absence of $BC$; and, also, there may be no position of minimum pressure, if the velocity be sufficiently great.

The same integral forms are obtained for $p$ between $A$ and $B$, $B$ and $C$ as in the separate cases; but whereas the constant $k$ had different values in these cases, it is easily seen (by reason of the continuity of velocity at $B$) that $k$ must have the same value for a given velocity in both sections. This value will be intermediate between the two already obtained.

Let $\varpi$ be the pressure at $B$, then for the motion between $A$ and $B$ we have the conditions $p = \Pi$, $x = 0$; $p = \varpi$, $x = 1$. Substituting these in equation (14) and subtracting, we obtain

$$\tfrac{1}{2}\log\frac{b\Pi^2 d_0^2/\mu - 6U\Pi d_0 + k}{b\varpi^2 d_1^2/\mu - 6U\varpi d_1 + k} - \log(d_0/d_1) + \frac{3\mu U}{K}\left[\tan^{-1}\frac{b\Pi d_0 - 3\mu U}{K} - \tan^{-1}\frac{b\varpi d_1 - 3\mu U}{K}\right] = 0 \quad \ldots\ldots(17).$$

Treating the equation for $p$ between $B$ and $C$ in the same way, we have

$$\tfrac{1}{2}\log\frac{b\Pi^2 d_0^2/\mu + 6U\Pi d_0 - k}{b\varpi^2 d_1^2/\mu + 6U\varpi d_1 - k} - \log(d_0/d_1)$$
$$+\frac{3\mu U}{2K'}\log\frac{(K' + b\Pi d_0 + 3\mu U)(K' - b\varpi d_1 - 3\mu U)}{(K' - b\Pi d_0 - 3\mu U)(K' + b\varpi d_1 + 3\mu U)} = 0 \ldots\ldots(18),$$

where $K' = \mu k b + 9U^2\mu^2$.

The equations (17) and (18) determine the value of the constants $k$, $\varpi$.

From the previous numerical calculation, if we take $d_0 = 10^{-4}$, $d_1 = \cdot 4 \times 10^{-4}$, $U = 5$, it is known that $\varpi = 7 \times 10^4$, $k = 129\cdot 2$ satisfy (17), and $\varpi = 7 \times 10^4$, $k = 111\cdot 3$ satisfy (18). Now assume $\varpi = 7\cdot 5 \times 10^4$ (say), and find the values of $k$ which satisfy (17) and (18), respectively.

These two sets of values can be plotted, and the intersection of the lines joining them gives approximately the common solutions of (17) and (18), some slight adjustment being necessary.

The following tables give the results of the calculations. The corresponding calculations in the case of incompressible liquid are not repeated, since they are the same as in the two separate cases, and have already been given.

$d_0 = 10^{-4}$, $d_1 = \cdot 4 \times 10^{-4}$

| $U$ | 5 | 10 | 20 |
|---|---|---|---|
| $(p_{\text{max.}} - \Pi)\,10^{-4}$ | ·98 | 2·08 | 3·98 |
| $(p_{\text{min.}} - \Pi)\,10^{-4}$ | − ·55 | − ·72 | − ·67 |
| $(\varpi - \Pi)\,10^{-4}$ | ·47 | 1·54 | 3·43 |
| $x_{\text{max.}}$ | ·80 | ·85 | ·90 |
| $x_{\text{min.}}$ | ·60 | ·49 | ·34 |

$d_0 = \cdot 8 \times 10^{-4}$, $d_1 = \cdot 6 \times 10^{-4}$

| $U$ | 5 | 10 | 20 |
|---|---|---|---|
| $(p_{\text{max.}} - \Pi)\,10^{-4}$ | ·28 | ·60 | 1·12 |
| $(p_{\text{min.}} - \Pi)\,10^{-4}$ | − ·15 | − ·20 | − ·18 |
| $(\varpi - \Pi)\,10^{-4}$ | ·12 | ·41 | ·95 |
| $x_{\text{max.}}$ | ·68 | ·75 | ·84 |
| $x_{\text{min.}}$ | ·47 | ·37 | ·24 |

It will be seen from these tables that the effect of placing together the two portions $AB$, $BC$ is to increase both the maximum and the minimum values of the pressure. It is clear from both tables that the minimum pressure decreases and increases again with increasing velocity, while its position approaches continuously to $C$; ultimately the pressure reaches a maximum at $B$ and then falls continuously to atmospheric pressure at $C$. The behaviour of the air, therefore, for large velocities is in marked contrast with that of an incompressible liquid.

# IV. *The Superior and Inferior Indices of Permutations.*

By Major P. A. MacMahon, F.R.S., Hon. Member Camb. Phil. Soc.

[*Received* 12 January 1914.]

Reference is made to the paper on Indices of Permutations*.

I here define indices of a new kind and subject them to investigation.

Let any assemblage of letters be

$$\alpha_1^{i_1}\alpha_2^{i_2}\ldots\alpha_s^{i_s}$$

and consider any permutation of them.

If any letter precedes $p_1'$ letters which have a smaller subscript we obtain the component $p_1'$ of the Superior Index.

The Superior Index of the permutation is defined to be

$$\Sigma p_1' = p',$$

the summation being in respect to every letter of the permutation.

On the other hand if any letter precedes $r_1'$ letters which have a larger subscript we obtain the component $r_1'$ of the Inferior Index.

The Inferior Index is defined to be

$$\Sigma r_1' = r',$$

the summation being in respect to every letter of the permutation.

Ex. gr. Consider the assemblage $\alpha^4\beta^3\gamma^2\delta$ and the permutation

$\beta\alpha\alpha\delta\gamma\alpha\alpha\beta\beta\gamma$

4 64 , $4+6+4=14$ the Superior Index

355 3311 , $3+5+5+3+3+1+1=21$ the Inferior Index.

If the permutation be reversed

$\gamma\beta\beta\alpha\alpha\gamma\delta\alpha\alpha\beta$

744 33 , $7+4+4+3+3=21$ the Superior Index

122331 11 , $1+2+2+3+3+1+1+1=14$ the Inferior Index,

and we see that the Superior and Inferior Indices of the permutation are respectively equal to

* "The Indices of Permutations and the Derivation therefrom of Functions of a Single Variable associated with the Permutations of any Assemblage of Objects." *American Journal of Mathematics*, Vol. xxxv. No. 3, 1913.

the Inferior and Superior Indices of the reversed permutation. This is obviously true in general so that we can assert in regard to the assemblage

$$\alpha_1^{i_1}\alpha_2^{i_2}\ldots\alpha_s^{i_s}$$

that the collection of numbers which specifies the superior indices of the permutations is identical with the collection which specifies the inferior indices.

Hence, in regard to the permutations of any assemblage,

$$\Sigma x^{p'} = \Sigma x^{r'}.$$

It is also readily established that, for every permutation,

$$p' + r' = \Sigma i_1 i_2;$$

for consider that part of a permutation which involves two letters $\alpha_k$, $\alpha_m$.

Suppose it to be

$$\ldots\alpha_k^{i_k'}\ldots\alpha_m^{i_m'}\ldots\alpha_m^{i_m''}\ldots\alpha_k^{i_k''}\ldots\alpha_k^{i_k'''}\ldots\alpha_m^{i_m'''}\ldots.$$

The portion of the superior index due to these two letters is, if $k > m$,

$$i_k'(i_m' + i_m'' + i_m''' + \ldots) + (i_k'' + i_k''')(i_m''' + \ldots)$$

and the portion of the inferior index is

$$(i_m' + i_m'')(i_k'' + i_k''' + \ldots) + i_m'''(\ldots).$$

Adding these together we find that the two letters contribute to the sums of the two indices the number

$$i_m i_k.$$

Thence obviously $$p' + r' = \Sigma i_1 i_2,$$

leading to the relation $$\Sigma x^{p'} = x^{\Sigma i_1 i_2}\Sigma x^{-p'}.$$

The maximum value of $p'$ is clearly $\Sigma i_1 i_2$ and its average value $\frac{1}{2}\Sigma i_1 i_2$.

The function $\Sigma x^{p'}$ is of degree $\Sigma i_1 i_2$ in $x$ and if it be divided by $x^{\frac{1}{2}\Sigma i_1 i_2}$ it is unaltered by the substitution of $\frac{1}{x}$ for $x$ because

$$\Sigma x^{p' - \frac{1}{2}\Sigma i_1 i_2} = \Sigma x^{\frac{1}{2}\Sigma i_1 i_2 - p'}.$$

A function of $x$ which satisfies these conditions is

$$\frac{(1-x)(1-x^2)\ldots\ldots(1-x^{i_1+i_2+\ldots+i_s})}{(1-x)(1-x^2)\ldots(1-x^{i_1}).(1-x)(1-x^2)\ldots(1-x^{i_2})\ldots\ldots(1-x)(1-x^2)\ldots(1-x^{i_s})}$$

and it will be shewn that this is in fact equal to $\Sigma x^{p'}$.

In the first place consider the assemblage $\alpha^i\beta^j$, and write $\Sigma x^{p'} = F_x(i, j)$.

All permutations which terminate with $\beta$, contribute

$$F_x(i, j-1)$$

to $F_x(i, j)$, and those which terminate with $\alpha$, contribute

$$x^j F(i-1, j).$$

Hence the difference equation

$$F_x(i, j) = x^j F_x(i-1, j) + F_x(i, j-1),$$

the solution of which, satisfying the above conditions, is

$$F_x(i, j) = \frac{(\mathbf{1})(\mathbf{2})\ldots(\mathbf{i}+\mathbf{j})}{(\mathbf{1})(\mathbf{2})\ldots(\mathbf{i}).(\mathbf{1})(\mathbf{2})\ldots(\mathbf{j})}$$

where $(\mathbf{m})$ has been written to denote (after Cayley) $1 - x^m$.

Similarly for the assemblage $\alpha^i\beta^j\gamma^k$, write $\Sigma x^{p'} = F_x(i, j, k)$. The permutations terminating with $\gamma$, $\beta$, $\alpha$ respectively contribute

$$F_x(i, j, k-1),\ x^k F_x(i, j-1, k) \text{ and } x^{j+k} F_x(i-1, j, k-1) \text{ to } F_x(i, j, k),$$

leading us to the difference equation

$$F_x(i, j, k) = x^{j+k} F_x(i-1, j, k) + x^k F_x(i, j-1, k) + F_x(i, j, k-1),$$

the solution of which, satisfying the conditions, is

$$F_x(i, j, k) = \frac{(\mathbf{1})\ldots(\mathbf{i}+\mathbf{j}+\mathbf{k})}{(\mathbf{1})\ldots(\mathbf{i}).(\mathbf{1})\ldots(\mathbf{j}).(\mathbf{1})\ldots(\mathbf{k})}.$$

Similarly we reach the difference equation

$$F_x(i_1, i_2, i_3 \ldots i_s) = x^{i_2+i_3+\ldots+i_s} F_x(i_1 - 1, i_2, \ldots i_s)$$
$$+ x^{i_3+\ldots+i_s} F_x(i_1, i_2 - 1, \ldots i_s) + \ldots + F_x(i_1, i_2, \ldots i_s - 1),$$

the solution of which, satisfying the conditions, is

$$F_x(i_1, i_2, \ldots i_s) = \frac{(\mathbf{1})\ldots(\mathbf{i}_1+\mathbf{i}_2+\ldots+\mathbf{i}_s)}{(\mathbf{1})\ldots(\mathbf{i}_1).(\mathbf{1})\ldots(\mathbf{i}_2)\ldots\ldots(\mathbf{1})\ldots(\mathbf{i}_s)}.$$

This result is remarkable because it establishes that

$$\Sigma x^{p'} = \Sigma x^p,$$

where $p$ is the Greater Index of a permutation (vide *American Journ. Math.* Vol. XXXV. No. 3, 1913). In fact the whole collection of Superior Indices coincides with the whole collection of Greater Indices, but it is not easy to establish this by a one-to-one correspondence.

Observe the permutations of $\alpha\alpha\beta\gamma$.

| Permutation | Greater Index | Superior Index |
|---|---|---|
| $\alpha\alpha\beta\gamma$ | 0 | 0 |
| $\alpha\alpha\gamma\beta$ | 3 | 1 |
| $\alpha\beta\alpha\gamma$ | 2 | 1 |
| $\alpha\gamma\alpha\beta$ | 2 | 2 |
| $\alpha\beta\gamma\alpha$ | 3 | 2 |
| $\alpha\gamma\beta\alpha$ | 5 | 3 |
| $\beta\alpha\alpha\gamma$ | 1 | 2 |
| $\gamma\alpha\alpha\beta$ | 1 | 3 |
| $\beta\alpha\gamma\alpha$ | 4 | 3 |
| $\gamma\alpha\beta\alpha$ | 4 | 4 |
| $\beta\gamma\alpha\alpha$ | 2 | 4 |
| $\gamma\beta\alpha\alpha$ | 3 | 5 |

$$\Sigma x^p = \Sigma x^{p'} = \frac{(\mathbf{1})(\mathbf{2})(\mathbf{3})(\mathbf{4})}{(\mathbf{1})(\mathbf{2}).(\mathbf{1}).(\mathbf{1})} = \frac{(\mathbf{3})(\mathbf{4})}{(\mathbf{1})^2}.$$

The particular result

$$\Sigma x^{p'} = \frac{(1)(2)\ldots(i+j)}{(1)\ldots(i).(1)\ldots(j)}$$

establishes that the permutations of the assemblage $\alpha^i\beta^j$ which have a superior (or greater) index equal to $p'$ are equinumerous with the partitions of the number $p'$ into parts, not exceeding $i$ in magnitude and not exceeding $j$ in number.

The property of $\Sigma x^{p'}$ that is before us leads to interesting relations between the functions $F_x(i, j)$.

Write the assemblage $\alpha^i\beta^j$ in the form

$$\alpha^{i-a}\beta^{j-b}.\alpha^a\beta^b$$

wherein $a$, $b$ are any two numbers, such that $a \not> i$, $b \not> j$.

It is to be shewn that

$$F_x(i, j) = \Sigma x^{(j-b)a} F_x(i-a, j-b) F_x(a, b)$$

wherein the summation is in respect of every composition $a$, $b$ of the constant number $a+b$. The number zero is not excluded so that if for instance $a+b=4$, the summation will be in respect of the compositions 40, 31, 22, 13, 04, it being understood, as above stated, that $a \not> i$, $b \not> j$.

It will be admitted that when the permutations admit of representation in the form

Some permutation of $\alpha^{i-a}\beta^{j-b}$ followed by some permutation of $\alpha^a\beta^b$,

the expression $x^{(j-b)a} F_x(i-a, j-b) F_x(a, b)$ denotes $\Sigma x^{p'}$ for the permutations in question. If we sum this expression for all values of $a$ and $b$ which give permutations involving $i+j-a-b$ letters followed by permutations involving $a+b$ letters we must arrive at the expression of $\Sigma x^{p'}$ for the whole of the permutations of $\alpha^i\beta^j$.

Hence $$F_x(i, j) = \Sigma x^{(j-b)a} F_x(i-a, j-b) F_x(a, b)$$

where $a+b=$ any constant number.

This interesting relation between the functions $x$ has a very interesting particular case.

If $a+b=\sigma$ a constant, we have

$$F_x(i, j) = x^{\sigma j} F_x(\sigma, 0) F_x(i-\sigma, j) + x^{(\sigma-1)(j-1)} F_x(\sigma-1, 1) F_x(i-\sigma+1, j-1) + \ldots + F_x(0, \sigma) F_x(i, j-\sigma).$$

Putting $\sigma = j$ we obtain

$$F_x(i, j) = x^{j^2} F_x(j, 0) F_x(i-j, j) + x^{(j-1)^2} F_x(j-1, 1) F_x(i-j+1, j-1) + \ldots + F_x(0, j) F_x(i, 0)$$

and if we now put $i=j$

$$F_x(j, j) = x^{j^2}\{F_x(j, 0)\}^2 + x^{(j-1)^2}\{F_x(j-1, 1)\}^2 + x^{(j-2)^2}\{F_x(j-2, 2)\}^2 + \ldots + \{F_x(0, j)\}^2.$$

This result is a generalization of the theorem in regard to the sum of the squares of the binomial coefficients, for putting $x=1$, it becomes

$$\binom{2j}{j} = \binom{j}{0}^2 + \binom{j}{1}^2 + \binom{j}{2}^2 + \ldots + \binom{j}{j}^2.$$

In general the reader will see that we have the relation

$$F_x(i_1 i_2 \dots i_s)$$
$$= \Sigma x^{(i_s - a_s)(a_1 + a_2 + \dots + a_{s-1}) + (i_{s-1} - a_{s-1})(a_1 + a_2 + \dots + a_{s-2}) + \dots + (i_2 - a_2) a_1}$$
$$. F_x(i_1 - a_1,\ i_2 - a_2, \dots i_s - a_s) . F_x(a_1 a_2 \dots a_s),$$

the summation being for every composition of a given number $a_1 + a_2 + \dots + a_s$ into $s$ or fewer parts $a_1, a_2, \dots a_s$ such that $a_s \ngtr i_s$ for all values of $s$.

The Superior Index as defined is obtained by adding several numbers together. This is the simplest way of obtaining the index, but the numbers so added are not the most interesting that come up for consideration. If $v > u$ the letter $\alpha_v$ adds a number to the index if it precedes one or more letters $\alpha_u$. Denote by $p'_{vu}$ the number added to the index due to the positions of the letters $\alpha_u$, $\alpha_v$. Moreover $\alpha_v$ may precede 1, 2, ... or $i_u$ letters $\alpha_u$. Denote by $p'_{vu,\sigma}$ the number of letters $\alpha_v$ which precede exactly $\sigma$ letters $\alpha_u$. Every time an $\alpha_v$ precedes exactly $\sigma$ letters $\alpha_u$ the number $\sigma$ is added to the index. Hence

$$p'_{vu} = p'_{vu,1} + 2p'_{vu,2} + 3p'_{vu,3} + \dots + i_u p'_{vu,i_u}.$$

Also if $p''_{vu}$ denotes, in regard to the whole of the permutations, the sum of the numbers added to the indices by reason of the relative positions of the letters $\alpha_u$, $\alpha_v$ and $p''_{vu,\sigma}$ the number of times in the whole of the permutations that a letter $\alpha_v$ precedes exactly $\sigma$ letters $\alpha_u$,

$$p''_{vu} = p''_{vu,1} + 2p''_{vu,2} + 3p''_{vu,3} + \dots + i_u p''_{vu,i_u}.$$

Now we know the value of $p''_{vu}$ from the following consideration. In any permutation consider merely the letters $\alpha_u$, $\alpha_v$. If $r'_{uv}$ denotes the number added to the Inferior Index by the relative positions of these letters we see that

$$p'_{vu} + r'_{uv} = i_u i_v,$$

for any one letter $\alpha_v$ contributes to the sum of the Superior and Inferior indices the number $i_u$ and therefore the total of $i_v$ letters $\alpha_v$ contributes the number $i_u i_v$. Hence the *average* value of $p'_{vu}$ in a permutation is $\frac{1}{2} i_u i_v$ and thence the number contributed to the Superior Indices of all of the permutations by the relative positions of $\alpha_u$ and $\alpha_v$ is

$$p''_{vu} = \tfrac{1}{2} i_u i_v \frac{(\Sigma i)!}{i_1!\, i_2! \dots i_s!}.$$

It will now be proved that $p''_{vu,\sigma}$

has a value which is independent of the number $\sigma$.

Consider the permutations of the assemblage

$$\alpha_1^{i_1} \alpha_2^{i_2} \dots \alpha_{u-1}^{i_{u-1}} \alpha_u^{i_u+1} \alpha_v^{i_v-1} \alpha_{v+1}^{i_{v+1}} \dots \alpha_s^{i_s},$$

which is derived from the original assemblage by adding an $\alpha_u$ and subtracting an $\alpha_v$.

In any permutation fix the attention upon the $i_u + 1$ letters $\alpha_u$. Call the one on the extreme right the last $\alpha_u$, the one nearest to it the last but one $\alpha_u$, the next one again the last but two $\alpha_u$ and so on. Now delete the last $\alpha_u$ but $\sigma$ from the permutation and substitute for it the letter $\alpha_v$. We have thus an $\alpha_v$ followed by $\sigma$ letters $\alpha_u$ and the assemblage is the original assemblage of letters. We thus construct a case of an $\alpha_v$ followed by exactly $\sigma$ letters $\alpha_u$ from every one of the

$$\frac{(\Sigma i)!}{i_1!\, i_2! \dots i_{u-1}!\, (i_u + 1)!\, (i_v - 1)!\, i_{v+1}! \dots i_s!}$$

permutations of the assemblage

$$\frac{\alpha_u}{\alpha_v}\alpha_1^{i_1}\alpha_2^{i_2}\ldots\alpha_s^{i_s}.$$

Thence
$$p''_{vu,\sigma} = \frac{i_v}{i_u+1}\frac{(\Sigma i)!}{i_1!\,i_2!\ldots i_s!},$$

a value which is independent of $\sigma$.

Therefore
$$p''_{vu,\,0} = p''_{vu,\,1} = \ldots = p''_{vu,\,i_u},$$

leading to
$$p''_{vu} = \binom{i_u+1}{2}p''_{vu,\sigma} = \tfrac{1}{2}i_u i_v \frac{(\Sigma i)!}{i_1!\,i_2!\ldots i_s!},$$

and
$$p''_{vu,\sigma} = \frac{i_v}{i_u+1}\frac{(\Sigma i)!}{i_1!\,i_2!\ldots i_s!}.$$

We deduce that the average value of $p''_{vu,\sigma}$ in a permutation is

$$\frac{i_v}{i_u+1}.$$

To illustrate these results take the assemblage $\alpha\alpha\beta\gamma$ wherein $i_1 = 2$, $i_2 = 1$, $i_3 = 1$.

| | $\beta\alpha$ | $\gamma\alpha$ | $\gamma\beta$ |
|---|---|---|---|
| $\alpha\alpha\beta\gamma$ | 0 | 0 | 0 |
| $\alpha\alpha\gamma\beta$ | 0 | 0 | 1 |
| $\alpha\beta\alpha\gamma$ | 1 | 0 | 0 |
| $\alpha\gamma\alpha\beta$ | 0 | 1 | 1 |
| $\alpha\beta\gamma\alpha$ | 1 | 1 | 0 |
| $\alpha\gamma\beta\alpha$ | 1 | 1 | 1 |
| $\beta\alpha\alpha\gamma$ | 2 | 0 | 0 |
| $\gamma\alpha\alpha\beta$ | 0 | 2 | 1 |
| $\beta\alpha\gamma\alpha$ | 2 | 1 | 0 |
| $\gamma\alpha\beta\alpha$ | 1 | 2 | 1 |
| $\beta\gamma\alpha\alpha$ | 2 | 2 | 0 |
| $\gamma\beta\alpha\alpha$ | 2 | 2 | 1 |

Here
$$p''_{21,\,0} = p''_{21,\,1} = p''_{21,\,2} = \frac{1}{3}\frac{4!}{2!\,1!\,1!} = 4,$$

and we verify that in the first column the numbers 0, 1 and 2 each occur 4 times.

Also
$$p''_{21} = \binom{3}{2}.4 = 12,$$

and we verify that the sum of the numbers in the first column is 12.

Again
$$p''_{31,\,0} = p''_{31,\,1} = p''_{31,\,2} = \frac{1}{3}.12 = 4,$$

$$p''_{31} = \binom{3}{2}.4 = 12,$$

and we verify that in the second column the numbers 0, 1 and 2 each occur 4 times and that the sum of the numbers is 12.

Again
$$p''_{32,\,0} = p''_{32,\,1} = \frac{1}{2}.12 = 6,$$

$$p''_{32} = \binom{2}{2}6 = 6,$$

and we verify that in the third column the numbers 0 and 1 each occur 6 times and that the sum of the numbers is 6.

## V. The Domains of Steady Motion for a Liquid Ellipsoid, and the Oscillations of the Jacobian Figure.

By R. HARGREAVES, M.A.

[*Received* 8 February 1914.]

ONE of the oscillations of ellipsoidal type for Maclaurin's figure of equilibrium has, at the junction with Jacobi's series, a period exactly one-half that of rotation; i.e. if a day means a period of rotation, the natural equatorial tide is here semi-daily.

This isolated result was reached some twenty years ago, and seemed of sufficient interest to stimulate enquiry into the course of the periods of oscillation of the Jacobian figure.

As the present work is of recent date the stimulus has been tardy in operation.

The scope of the investigation has been extended to cover other matters which, like the question of periods of oscillation, require for their complete discussion much laborious calculation with transcendental equations. The results are made accessible by the use of diagrams to represent the domains of steady motion for a homogeneous liquid ellipsoid under its own gravitation, and an inspection of these is sufficient to shew what kinds of steady motion are possible for an ellipsoid of given shape.

Special attention is given to the Jacobian form where a full series is treated with reference to shape, angular velocity and momentum, and kinetic energy; while the periods of the ellipsoidal oscillations are added for a smaller number of cases, sufficient to make the course clear through the entire range. In respect to the Jacobian an interesting feature is the connexion of the movement in values of angular velocity and momentum along the series, with the quantities on which secular stability depends.

With respect to motion about two axes the most interesting point is that the conditions laid down by Riemann for his Case II are entirely superseded by the condition of positive pressure.

It is proposed to describe the main results in general terms before proceeding to the analysis on which they are based.

For the spheroids the oscillations may be called *polar* and *equatorial*; in the former the equator remains a circle but its radius and the polar axis are subject to periodic change, in the latter the polar axis is unaltered, the equator suffers a periodic elliptical deformation.

The oscillations of the Jacobian near the opening of the series differ little from those of the spheroid; as the form moves away from the spheroidal the terms polar and equatorial fail to describe them, but in the ultimate position the oscillations become respectively equatorial and polar if the word polar is now applied to the long axis, equator to the nearly circular ellipse containing the short axis.

At the initial point of Jacobi's series as stated above the equatorial oscillation has exactly a semi-daily period, the polar oscillation has a slightly shorter period. The periods diverge as the ellipsoid is elongated, the former increasing the latter decreasing. In the limit of extreme elongation the former period is daily, the latter has a finite value while the period of rotation is indefinitely long. For a certain range from the initial Jacobian, and also for a range of Maclaurin's spheroids on each side of the junction, the periods differ little from the half-day in excess or defect, both cases being represented.

The position for spheroids is that the frequencies $n_e$ and $n_p$ are finite for the spherical form where the rotation is indefinitely slow; the former, at first the greater, falls the more rapidly with increase of oblateness, and equality is attained for

$$c/a = \cdot 5892 = \cos 53^\circ\, 54', \quad \text{when} \quad n^2/\omega^2 = 4\tfrac{1}{6}.$$

The Jacobian junction is reached when

$$c/a = \cdot 5827 = \cos 54^\circ\, 21'\, 27'', \quad \text{and} \quad n_e = 2\omega, \quad n_p^2/\omega^2 = 4\cdot 1182.$$

The value $n_p = 2\omega$ is reached for

$$c/a = \cdot 5612 = \cos 55^\circ\, 52'.$$

With these values may be compared the entries in the table for Jacobians for values of $\alpha$ up to 30°.

Now any external body in the presence of which a liquid ellipsoid is rotating, will shew a period something more or less than the day according as the relative orbital motion is direct or retrograde, and its quasi-statical tidal influence will have a period near the half-day. Accordingly ellipsoids of a shape deviating to some moderate extent from that at the Maclaurin-Jacobi junction will be specially sensitive to the tides induced, in consequence of the closeness of periods of the natural and forced oscillations.

The first calculations of periods were based on material provided in Sir George Darwin's paper on 'Jacobi's Figure of Equilibrium*.' Some irregularity appeared in the succession of values for these periods, and a new series of points was determined, generally in close agreement with Darwin's paper. In one point of some importance there is disagreement. Darwin found a maximum for the velocity of rotation at some little distance from the beginning of the Jacobian series, I find an uninterrupted fall. This question and that of the rise in value of the momentum appear to be connected with quantities occurring in what is called the test of secular stability, as restricted to deformations consistent with ellipsoidal shape. I find this restricted test to be satisfied through the whole range of Jacobians, and in connexion with it a regular fall in velocity of rotation, a rise in angular momentum, and for kinetic energy a rise to a maximum situated, as Darwin found it, in the range where elongation is considerable.

Ellipsoids of given shape may be represented on a diagram by coordinates $x$, $y$ the ratios of one axis to the two others. If we take a standard order $a > b > c$ and write $x = c/a$, $y = c/b$, then with $x \leqslant 1$, $y \leqslant 1$, $x \leqslant y$ the representative point lies in a triangle

* *Proceedings Royal Society*, 1886, pp. 319—336. Until after the communication of this paper the author was not aware that Darwin's paper had undergone a thorough revision in preparation for the issue of his *Collected Papers*. The corrections there made have removed the discrepancies to which reference is made here and in § 12.

*OSP*. *OS* gives Maclaurin's series, *O* the disk end, *S* the spherical end, *SP* represents prolate spheroids with *P* for the case of extreme elongation. For points near *S*, *P*, *O* the direction of deviation from these positions is important and we find boundaries different in character reaching these points as a terminus in different directions. Thus points near *O* make $c/a$ and $c/b$ both small with ratio $y:x = a:b$ finite so that elliptical disks are represented, with a circular disk for $y = x$. The line *OP* and its neighbourhood represent the series of cases in which with $c$ always small, the ratio $b:a$ has degrees of smallness falling to the case $b = c$ at *P*.

For one case, that of rotation about a single axis, I have found it more convenient to take the axis of rotation as numerator whether it is least or mean axis. In this way we separate the fields of rotation about mean and least axis, which are dynamically distinct, leaving them contiguous along the line *SP* at which one merges into the other.

The standard system has been used in all other diagrams and has the advantage of shewing by the overlapping of areas what shapes are capable of more than one type of steady motion.

The boundaries of the field of steady motion about one axis are determined by the vanishing of Riemann's constants $\tau$ and $\tau'$. The line *RP* (fig. 1) represents $\tau = 0$, and

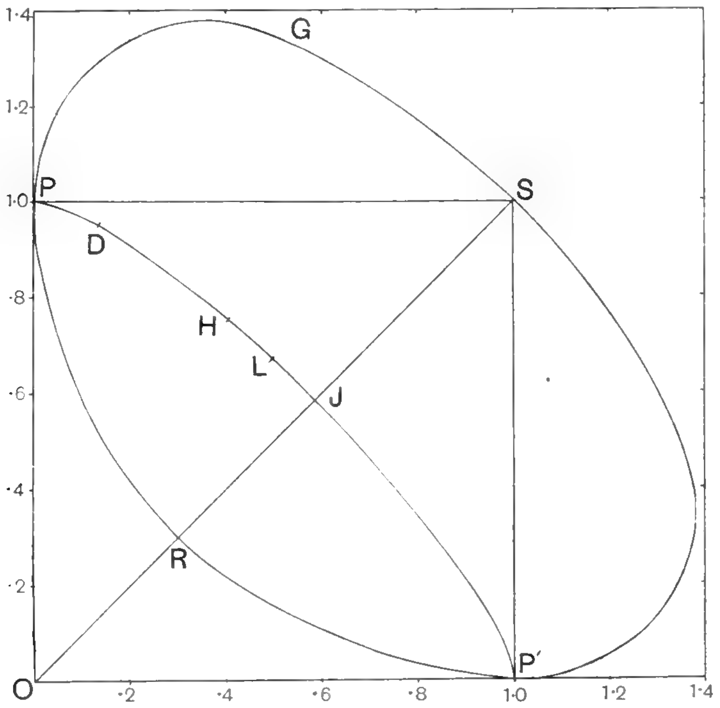



Fig. 1. Field of steady motion about one axis.

rotation about the least axis covers the field *RSP*. Along the line *SP* it merges in the case of rotation about mean axis the field for which extends to the boundary $\tau' = 0$ or *SGP*. The figure may be duplicated by interchanging $x$ and $y$ or taking the two axes about which there is no rotation in a different order. The whole field is then a flattened oval, with Maclaurin's line for a line of symmetry, while the Jacobian is a curved diagonal from $P$ where one axis is indefinitely increased to $P'$ where the other is increased, and cuts $OS$ at right angles. This affords a more satisfactory representation of the relations of these lines to the field than we get by taking half the figure and regarding Maclaurin's line as a third boundary. [At the same time $\tau = 0$ for Maclaurin's case, vanishing through the factor $(a-b)^2$ while for the curve $RP$ the vanishing factor is transcendental.]

The point $R$ where Riemann found instability for the spheroids is the point where $OS$ ceases to be in contact with the general field. Here $n_e^2$ vanishes and becomes negative; $n_p^2$ is not affected because the polar oscillation does not postulate a difference between $a$ and $b$, it remains positive and is ultimately $=\omega^2$.

If we use the least axis as numerator of coordinates the field of motion about a mean axis is represented by a space below $SP$. The boundary $SHP$ (fig. 2) proceeds from $S$ along

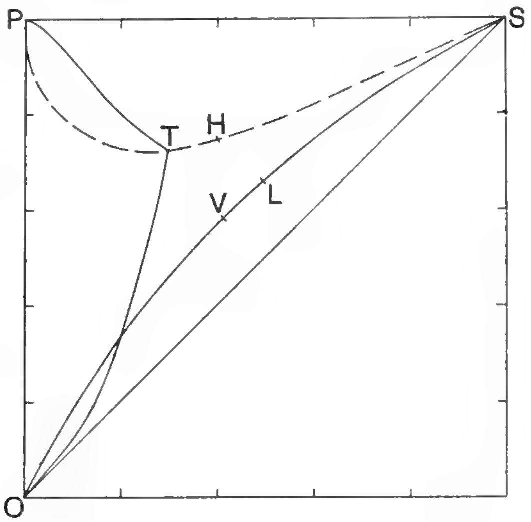



Fig. 2. Field of steady motion about two axes. Riemann's Cases I and III.

the line $y+1=2x$, crosses the Jacobian at $H$ and reaches $P$ touching $OP$. For Jacobians between $H$ and $P$ an alternative motion is possible having the mean axis as axis of rotation. At $H$ the values of the elements in these alternative motions are different; the point $H$ cannot be a point of bifurcation in the usual sense, and there is no ground for supposing it to be a starting point of possible instability in the Jacobian series. The algebraical inequalities which are known to exist in connexion with the transcendental curves $\tau$, $\tau'$, $J$, are not in fact of much service in drawing the curves: they are generally wide except near a terminal point. But the condition

$$x^2+xy+y^2<1 \text{ for } \tau,$$

when reversed gives some clue to the Jacobian line. The inequality $x+y<2x_0$ is close for a considerable range, and in these coordinates the principal section of a flat saucer is a good representation of the line.

There are three cases of steady motion involving two axes, for which we shall use the standard order $a>b>c$ and coordinates $x=c/a$, $y=c/b$.

Case I, in which the *greatest* and *least* axes are those of rotation, has for field a narrow lune (fig. 2) skirting the line $OS$ and bounded by the hyperbola $y(1+x)=2x$ representing the equality $2b=a+c$. A vertex $V$ of the hyperbola ($x=\sqrt{2}-1$, $y=2-\sqrt{2}$) lies symmetrically with regard to $OS$, the boundary has at $S$ a tangent $x+1=2y$ in common with the boundary for rotation about a mean axis, but the fields lie on opposite sides.

Case II, in which again the *greatest* and *least* axes are those of rotation, is defined in a preliminary way by $OP$, an arc $OQ$ (fig. 3) of the hyperbola $y(1-x)=2x$ representing $2b=a-c$, and an arc $QP$ of the quartic $y^2-4x^2=y^2(y^2-x^2)$. The intersection

$$x=2-\sqrt{3}, \quad y=\sqrt{3}-1$$

is the point where the loop of the quartic has $x$ a maximum.

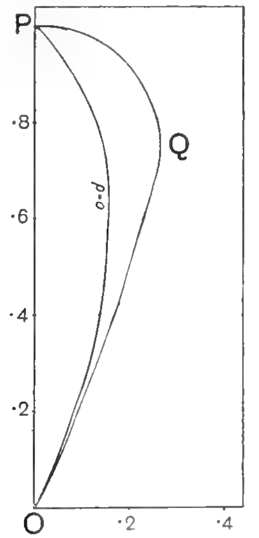



Fig. 3. Field of steady motion about two axes. Riemann's Case II.

In this connexion an interesting feature emerges. For all other cases of steady motion the condition for a positive value of the pressure is satisfied as a necessary consequence of other conditions. In this one case it is not; it proves to be more stringent than either algebraical relation, and the line of zero pressure therefore supersedes them as the effective boundary. If we suppose motion possible between the pressure line and the algebraical boundaries, this is the one region in the whole range of steady motions where there is a manifest occasion for the separation into distinct masses.

Case III, in which the *greatest* and *mean* axes are those of rotation, has a field $OTP$ (fig. 2).

The part $OT$ is the hyperbola $y(1-2x)=x$ representing $2c=a-b$, the part $TP$ is given by a transcendental equation.

It will be noticed that the boundary of I is single and algebraical, that of II single and transcendental, that of III composite with sections of each type.

With the aid of the diagrams it is easy to examine all the cases where overlapping shews that more than one type of motion is possible for certain shapes of ellipsoid. We may note that there is a small region in the neighbourhood of $x = \cdot 22$, $y = \cdot 38$, not a case of extreme dimensions, for which no state of steady motion exists. [*PR* of fig. 1 crosses *OT* and *OV* of fig. 2 above their point of intersection.]

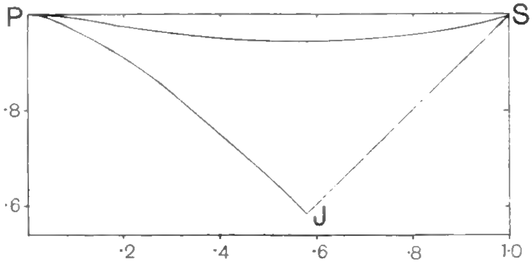



Fig. 4. Field of Roche's steady motion. Contour shews extreme cases.

Roche has made use of equilibrium forms* in which the attraction of a distant body is taken into account, with the limitation that the rotation and relative orbital revolution have the same period. The connexion with the above seems sufficient to justify the inclusion of this case in the diagrams.

The case represented is that in which the liquid mass is extremely small in relation to the attracting body, and the interval between the smooth curve *SP* and the broken line *SJP* corresponding to the other extreme where the distant body is small, will be bridged by a

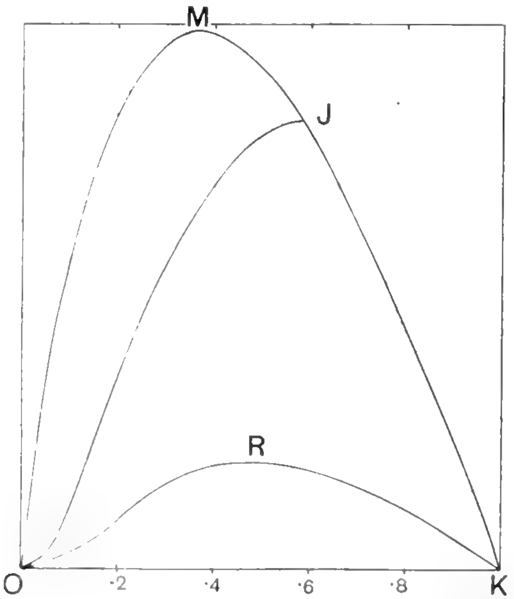



Fig. 5. Graph of $\omega^2$ for figures of Maclaurin, Jacobi and Roche.

succession of intermediate forms. The graph for $\omega^2/4\pi\rho$ is also set in the diagram (fig. 5) for the same quantity in relation to Jacobi's ellipsoids and Maclaurin's spheroids.

* I have only seen the reference in M. Poincaré's *Hypothèses Cosmogoniques*, p. 54.

§ 1. Riemann's* equations to determine the frequencies $n$ and the ratios of amplitudes $da:db:dc$ may be written

$$\left.\begin{aligned}\{a^2E_{aa}-n^2(a^2+c^2)\}\frac{da}{a}+(abE_{ab}-n^2c^2)\frac{db}{b}=0\\ \{b^2E_{bb}-n^2(b^2+c^2)\}\frac{db}{b}+(abE_{ab}-n^2c^2)\frac{da}{a}=0\end{aligned}\right\}\quad\ldots\ldots(1),$$

giving a quadratic in $n^2$ when the ratios are eliminated. The total energy is $mE/5$, and $E_{aa}$ denotes the second differential coefficient of $E$ with regard to $a$, when $a$ and $b$ are independent variables with which $c$ is connected by the condition of constant volume.

The steady motion for which the oscillation is considered is given by

$$\left.\begin{aligned}E_a\equiv\frac{Aa^2-Cc^2}{a}-\frac{2\tau^2}{(a-b)^3}-\frac{2\tau'^2}{(a+b)^3}=0\\ E_b\equiv\frac{Bb^2-Cc^2}{b}+\frac{2\tau^2}{(a-b)^3}-\frac{2\tau'^2}{(a+b)^3}=0\end{aligned}\right\}\quad\ldots\ldots(2),$$

in which $\tau$ and $\tau'$ are Riemann's constants of integration, and

$$A=2\pi\rho abc\int_0^\infty d\lambda/\sqrt{(a^2+\lambda)^3(b^2+\lambda)(c^2+\lambda)}.$$

Differentiating (2) we get $E_{aa}$ and substitute the values of $\tau^2$, $\tau'^2$ from (2), so that the final expressions are in terms of $a$, $b$, $c$ only.

It is convenient to use integrals

$$F,G,H\equiv 2\pi\rho abc\int_0^\infty\frac{(1,\lambda,\lambda^2)\,d\lambda}{\{(a^2+\lambda)(b^2+\lambda)(c^2+\lambda)\}^{\frac{3}{2}}}\quad\ldots\ldots(3).$$

When $a$, $b$, $c$ are treated as three independent variables,

$$b\frac{\partial A}{\partial b}=Gc^2+H$$

and similar equations, while $A+B+C=4\pi\rho$ gives

$$a\frac{\partial A}{\partial a}+G(b^2+c^2)+2H=0,$$

and $aE_a$ is $a\dfrac{\partial E}{\partial a}-c\dfrac{\partial E}{\partial c}$. By this method we derive from (2),

$$a^2E_{aa}=4Fa^2b^2c^2+G(2a^2c^2+b^2c^2-a^2b^2)-2Ha^2+6a^2\left\{\frac{\tau^2}{(a-b)^4}+\frac{\tau'^2}{(a+b)^4}\right\},$$

$$abE_{ab}=2Fa^2b^2c^2+G(a^2c^2+b^2c^2-a^2b^2)-Hc^2-6ab\left\{\frac{\tau^2}{(a-b)^4}-\frac{\tau'^2}{(a+b)^4}\right\},$$

or on clearing $\tau^2$ and $\tau'^2$

$$\left.\begin{aligned}a^2E_{aa}=4Fa^2b^2c^2+G(5a^2c^2+b^2c^2-a^2b^2)+Ha^2\\ abE_{ab}=2Fa^2b^2c^2-2G(a^2c^2+b^2c^2-a^2b^2)-4Hc^2\end{aligned}\right\}\quad\ldots\ldots(4).$$

The quantities $A$, $B$, $C$ are connected with $F$, $G$, $H$ by

$$A=Fb^2c^2+G(b^2+c^2)+H\quad\ldots\ldots(5);$$

* Riemann, *Abh. K. Ges. Wiss. Göttingen*, vol. IX. 1860; *Collected Works*, p. 168 sqq.

and the relation

$$4\pi\rho = F\Sigma a^2b^2 + 2G\Sigma a^2 + 3H \qquad \text{.........(6)}$$

represents

$$A + B + C = 4\pi\rho.$$

These results apply to the whole domain of rotation about a single axis.

§ 2. For Jacobi's ellipsoid $2\tau/(a-b)^2 = 2\tau'/(a+b)^2 = \omega$, and in terms of $\omega$ equations (2) are

$$\omega^2 = (Aa^2 - Cc^2)/a^2 = (Bb^2 - Cc^2)/b^2 \qquad \text{.........(7)},$$

i.e. $\omega^2$ is given in terms of the ratios of axes, and these ratios have a connexion in virtue of which the series is represented by a line in a plane diagram.

The value of $\omega^2$ in terms of $G$, $H$ is

$$\omega^2 = Gc^2 + H \qquad \text{.........(8)},$$

and the equation of condition is

$$G(a^2c^2 + b^2c^2 - a^2b^2) + Hc^2 = 0 \qquad \text{.........(9)},$$

which somewhat simplifies (4), making

$$\left.\begin{aligned} a^2E_{aa} &= 4Fa^2b^2c^2 + 4Ga^2c^2 + H(a^2 - c^2) \\ abE_{ab} &= 2Fa^2b^2c^2 - 2Hc^2 \end{aligned}\right\} \qquad \text{.........(10)}.$$

The treatment of oscillations turns on the calculation of $F$, $G$, $H$. Readers of Darwin's paper will recall that $\omega^2/4\pi\rho$ is evaluated for a series of cases in which the geometrical condition is satisfied. Equations (6), (8) and (9) are then sufficient to determine $F$, $G$, $H$, viz.

$$F\Sigma a^2b^2 = 4\pi\rho - \omega^2\left\{1 + \frac{2a^2b^2 + c^4}{(a^2 - c^2)(b^2 - c^2)}\right\}, \quad G = \frac{\omega^2c^2}{(a^2 - c^2)(b^2 - c^2)}, \quad H = \omega^2\left\{1 - \frac{c^4}{(a^2 - c^2)(b^2 - c^2)}\right\}$$

$$\text{.........(11)}.$$

This is a convenient method of utilizing the material of Darwin's paper for the further purpose of dealing with oscillations, and, so far as I can see, his equations are the most convenient way of expressing the geometrical condition and determining $\omega^2$, *except in the neighbourhood of each end of the series.* As a test of accuracy the relation

$$3Fa^2b^2c^2 + 2G\Sigma a^2b^2 + H\Sigma a^2 = 4\pi\rho abc F(\alpha, \gamma)/\sqrt{a^2 - c^2} \qquad \text{.........(12)}$$

has been used; each member is an expression for

$$Aa^2 + Bb^2 + Cc^2.$$

Darwin expresses $A$, $B$, $C$ in terms of angles $\alpha$, $\beta$, $\gamma$, where

$$b = a\cos\beta, \quad c = a\cos\gamma, \quad \text{and } \sin\beta = \sin\alpha\sin\gamma \qquad \text{.........(13)};$$

and these appear in elliptic integrals $F(\alpha, \gamma)$ and $E(\alpha, \gamma)$, i.e.

$$E(\alpha, \gamma) = \int_0^\gamma \sqrt{1 - \sin^2\alpha\sin^2\gamma}\, d\gamma.$$

With a multiplier $\qquad M = \cos^2\alpha\sin^2\beta\sin\gamma/\cos\beta\cos\gamma,$

we have

$$\frac{AM}{4\pi\rho} = (F - E)\cos^2\alpha, \quad \frac{BM}{4\pi\rho} = E - F\cos^2\alpha - \frac{\sin^2\alpha\sin\gamma\cos\gamma}{\cos\beta}, \quad \frac{CM}{4\pi\rho} = \sin^2\alpha\left\{-E + \frac{\cos\beta\sin\gamma}{\cos\gamma}\right\}$$

$$\text{.........(14)}.$$

The equation of condition is

$$E(1+\sin^2\alpha\tan^2\beta\cos^2\gamma)=(2F-E)\cos^2\alpha+\sin\alpha\tan\beta\cos\gamma\,(1+\sin^2\beta)\ldots\ldots\ldots(15),$$

and

$$\frac{\omega^2}{4\pi\rho\cos\beta\cos\gamma}=\frac{F-E}{\sin^2\beta\sin\gamma}+\frac{E\cos^2\gamma}{\cos^2\alpha\sin^3\gamma}-\frac{\cos\beta\cos\gamma}{\cos^2\alpha\sin^2\gamma}\ldots\ldots\ldots\ldots(16).$$

§ 3. We proceed to examine the question of increase or decrease in angular velocity and momentum and in kinetic energy as we follow the Jacobian series from the junction.

This is closely connected with the variation which has been used as a test of secular stability, viz. that of

$$T+U\equiv\frac{5h^2}{2m(a^2+b^2)}-\frac{2\pi\rho mabc}{5}\int_0^\infty\frac{d\lambda}{\sqrt{(a^2+\lambda)(b^2+\lambda)(c^2+\lambda)}}\ldots\ldots\ldots\ldots\ldots(17),$$

subject to the condition of constant volume. Here use $\alpha\beta\gamma$ for $a^2b^2c^2$, $g$ for $a^2+b^2$, and write $g(1+g')$, $\gamma(1+\gamma')$ for values after variation. The alteration in $P$ or

$$(\alpha+\lambda)(\beta+\lambda)(\gamma+\lambda)$$

is to the second order

$$\lambda\left[gg'(\gamma+\lambda)+\gamma'\{\gamma(g+\lambda)-\alpha\beta\}\right]+\lambda\gamma'(g\gamma g'+\alpha\beta\gamma').$$

Thus with

$$h=\frac{m}{5}(a^2+b^2)\,\omega,$$

$$\left.\begin{aligned}\Delta_1\frac{10}{m}(T+U)&=-\omega^2gg'+2\pi\rho abc\int_0^\infty\frac{\lambda d\lambda}{P^{\frac32}}\left[gg'(\gamma+\lambda)+\gamma'\{\gamma(g+\lambda)-\alpha\beta\}\right]\\ \Delta_2\frac{10}{m}(T+U)&=\omega^2gg'^2+2\pi\rho abc\int_0^\infty\frac{\lambda d\lambda}{P^{\frac32}}(g\gamma g'+\alpha\beta\gamma')\gamma'\\ &\quad-\frac{3\pi\rho abc}{2}\int_0^\infty\frac{\lambda^2d\lambda}{P^{\frac52}}\left[gg'(\gamma+\lambda)+\gamma'\{\gamma(g+\lambda)-\alpha\beta\}\right]^2\end{aligned}\right\}\ldots(18).$$

Equating to zero the coefficients of $g'$ and $\gamma'$ in $\Delta_1$ we have the usual Jacobian relations

$$\omega^2=2\pi\rho abc\int_0^\infty\frac{\lambda(\gamma+\lambda)\,d\lambda}{P^{\frac32}},\quad 0=\int_0^\infty\frac{\lambda\{\gamma(g+\lambda)-\alpha\beta\}\,d\lambda}{P^{\frac32}}.$$

If the second variation is written

$$\left.\begin{aligned}&Lg'^2+M\gamma'^2+2Ng'\gamma',\\ &\frac{L}{\pi\rho abc}=2g\int_0^\infty\frac{\lambda(\gamma+\lambda)\,d\lambda}{P^{\frac32}}-\frac{3g^2}{2}\int_0^\infty\frac{\lambda^2(\gamma+\lambda)^2\,d\lambda}{P^{\frac52}}\\ &\frac{M}{\pi\rho abc}=2\alpha\beta\int_0^\infty\frac{\lambda d\lambda}{P^{\frac32}}-\frac32\int_0^\infty\frac{\lambda^2\{\gamma(g+\lambda)-\alpha\beta\}^2\,d\lambda}{P^{\frac52}}\\ &\frac{N}{\pi\rho abc}=g\gamma\int_0^\infty\frac{\lambda d\lambda}{P^{\frac32}}-\frac{3g}{2}\int_0^\infty\frac{\lambda^2(\gamma+\lambda)\{\gamma(g+\lambda)-\alpha\beta\}\,d\lambda}{P^{\frac52}}\end{aligned}\right\}\ldots\ldots\ldots\ldots\ldots(19).$$

Referring the integrals to a common denominator we find $L$, $M$ and $M-N$ to be positive, for the coefficient of each power of $\lambda$ in the numerator is positive when account is taken of $\alpha\beta>\gamma g$.

An equivalent* of the integral $0 = \int_0^\infty \frac{d}{d\lambda}\left\{\frac{\lambda^2(\lambda+\gamma)}{P^{\frac{3}{2}}}\right\} d\lambda$ is

$$\left.\begin{aligned} &\int_0^\infty \frac{\lambda(3\lambda - 4\gamma)\,d\lambda}{P^{\frac{3}{2}}} + \int_0^\infty \frac{3\lambda^2(\lambda+\gamma)}{P^{\frac{5}{2}}}\{2(\lambda\gamma + g\gamma - \alpha\beta) - g(\lambda+\gamma)\} = 0; \\ &\text{in virtue of which} \qquad \frac{N}{\pi\rho abc} = \frac{3g}{4}\int_0^\infty \frac{\lambda^2(\lambda+\gamma)(\lambda^2+\alpha\beta - g\gamma)\,d\lambda}{P^{\frac{5}{2}}}, \\ &\text{and then} \qquad \frac{L-2N}{\pi\rho abc} = \frac{g}{2}\int_0^\infty \frac{\lambda(\lambda+4\gamma)\,d\lambda}{P^{\frac{3}{2}}} \end{aligned}\right\} \quad \text{......(20).}$$

Thus $N$ also is positive and $L > 2N$, $M > N$. Since $L$, $M$ and $LM - N^2$ are positive the energy-expression (17) is a minimum for the Jacobian figure.

§ 4. The relation between $g'$ and $\gamma'$ which corresponds to movement along the Jacobian line is got by taking the variation of the equation of condition and is

$$0 = \int_0^\infty \frac{\lambda d\lambda}{P^{\frac{3}{2}}}\{\gamma\gamma'(g+\lambda) + \alpha\beta\gamma' + g\gamma g'\}$$

$$-\frac{3}{2}\int_0^\infty \frac{\lambda^2 d\lambda}{P^{\frac{5}{2}}}\{\gamma(g+\lambda) - \alpha\beta\}[gg'(\gamma+\lambda) + \gamma'\{\gamma(g+\lambda) - \alpha\beta\}],$$

or $$0 = \int_0^\infty \frac{\lambda d\lambda}{P^{\frac{3}{2}}}(g\gamma g' + 2\alpha\beta\gamma') - \frac{3}{2}\int_0^\infty \quad \text{,,} \qquad , \qquad \text{,,} \qquad \text{,,}$$

as modified by the equation of condition, i.e. the relation is $M\gamma' + Ng' = 0$.

The variation of $\omega^2$ is given by $\Delta\omega^2 = -g'Q + 2N\gamma'/g$ where

$$Q = 3g\int_0^\infty \frac{\lambda^2(\gamma+\lambda)^2\,d\lambda}{P^{\frac{5}{2}}},$$

so that $L/g = \omega^2 - Q/2$. The variation along the Jacobian line is

$$\Delta\omega^2 = -g'\left(Q + \frac{2N^2}{Mg}\right),$$

or if we prefer,

$$\frac{d(\omega^2)}{dg} = -\frac{1}{g}\left(Q + \frac{2N^2}{Mg}\right) = \frac{2}{g^2}\left(L - \frac{N^2}{M} - \omega^2 g\right) \quad \text{........................(21).}$$

As $M$ and $Q$ are both positive, $\omega^2$ falls continuously as $g$ increases, i.e. in passing along the Jacobian line from the junction. From the second form of (21) we get

$$\frac{d}{dg}(\omega^2 g^2) = 2\left(L - \frac{N^2}{M}\right) \quad \text{....................................(22).}$$

Also for this particular variation along the Jacobian line

$$\Delta_2 \frac{10}{m}(T+U) = g'^2\left(L - \frac{N^2}{M}\right), \quad \text{or} \quad \Delta_2 2(T+U) = \frac{\omega\Delta g\Delta h}{g},$$

using $\Delta h = \frac{m}{5}\Delta(g\omega)$. It is clear then that increase of energy, and of angular momentum both follow from the condition $LM > N^2$. But

$$\frac{d}{dg}(\omega^2 g) = \omega^2 - \left(Q + \frac{2N^2}{Mg}\right) = \frac{1}{g}\left\{2\left(L - \frac{N^2}{M}\right) - \omega^2 g\right\} \quad \text{.....................(23),}$$

the sign of which is not determined by the inequalities stated. The sign is at first positive

* This step occurs in C. O. Meyer's paper, *Crelle*, t. XXIV. (1842), where the continuous fall of $\omega^2$ is established.

but ultimately negative, the kinetic energy having a maximum for a somewhat elongated Jacobian (*v. infra* p. 76).

§ 5. The transition to the spheroid is interesting. With $a(1+a')$, $a(1+b')$ for the varied axes,

$$g' = a' + b' + \tfrac{1}{2}(a'^2 + b'^2), \text{ and } \gamma' = -1 + 1/(1+a')^2(1+b')^2$$

or

$$\gamma' = -2(a'+b') + 3(a'^2+b'^2) + 4a'b' = -2g' + 3g'^2 + (a'-b')^2.$$

In this case the independence of $\gamma'$ and $g'$ is only realised in terms of the second order. The part of the variation (18) which remains of first order is

$$2a^2 g' \left[ -\omega^2 + 2\pi\rho(a^2 - c^2)\, c \int_0^\infty \frac{\lambda(a^2+\lambda)\,d\lambda}{P^{\frac{3}{2}}} \right],$$

giving the value of $\omega^2$ appropriate to the spheroid. This value of $\omega^2$ appears in the second line of (18), and then the variation of second order in which now

$$\gamma' = -2g' = -2(a'+b'),$$

takes the form

$$(L + 4M - 4N)\, g'^2 + 2\pi\rho a^2 c \int_0^\infty \frac{\lambda\,d\lambda}{P^{\frac{3}{2}}} \{c^2(2a^2+\lambda) - a^4\}\{g'^2 + (a'-b')^2\}$$

or on reduction

$$2\pi\rho a^2 c\,(a'+b')^2 \int_0^\infty \frac{\lambda\,d\lambda}{(a^2+\lambda)^3(c^2+\lambda)^{\frac{1}{2}}} \left[ 2a^2 + c^2 + \frac{4c^2(a^2-c^2)}{c^2+\lambda} + \frac{3c^2(a^2-c^2)^2}{(c^2+\lambda)^2} \right]$$

$$+ 2\pi\rho a^2 c\,(a'-b')^2 \int_0^\infty \frac{\lambda\,d\lambda}{P^{\frac{3}{2}}} \{c^2(2a^2+\lambda) - a^4\},$$

a result verified by direct treatment of Maclaurin's case. The first term is always positive, the second positive from the spherical end to the Jacobian junction and thenceforward negative. The term which changes sign is a term carried over from the original first variation to what in terms of $a'$ and $b'$ is second variation, and the necessity for change is due to the circumstance that $a = b$ gives a relation between $g'$ and $\gamma'$ to the first order as the expression of the condition of constant volume.

§ 6. If we use for kinetic energy $\frac{m}{5}\left\{\frac{\tau^2}{(a-b)^2} + \frac{\tau'^2}{(a+b)^2}\right\}$ and after forming the variation write $\tau/(a-b)^2 = \tau'/(a+b)^2 = \omega/2$, the first variation agrees with (18), the second shews a positive increment, and the minimum property is more easily assured than with the above formula. The greater stringency of the dynamical minimum condition is in the case of the oblate spheroid represented by the withdrawal of the limit of validity from Riemann's point to Jacobi's junction.

The minimum theorem may be stated as follows: *a homogeneous body with kinetic energy due to rotation about a principal axis, and potential energy due to its own gravitation, will, when restricted to constant volume and ellipsoidal form, shew a minimum of total energy, if for values of the momentum below a certain limit the ellipsoid is an oblate spheroid, and for values above the limit it has the Jacobian form, and in each case the velocity is that given by hydrodynamical theory.*

The juxtaposition of kinetic and potential energy in the minimum problem postulates some mobility or capacity of accommodating shape to stress, but not the complete mobility of a liquid, for that demands Riemann's variation method; or if such mobility exists, then a frictional action is postulated sufficient to suppress, as they arise, all motions involving departure from that of a rigid body.

In pure hydrodynamics an existence theorem is established, the argument of secular stability extends the scope of its application.

If an external body is taken into account the problem is more complex, but it seems a probable forecast that the tendency will be to transfer angular momentum and with it energy, from rotation to relative orbital revolution; the historical order suggested is a passage from the direction of $P$ towards the Jacobian junction, and thence to the spheroidal form.

§ 7. We proceed to methods of approximation which it is advisable to use near each end of the series. Near the Maclaurin junction Jacobi's figure approximates to an oblate spheroid, at the other extreme to an elongated prolate spheroid. We can in these cases avoid the elliptic functions, and calculate $F$, $G$, $H$ directly.

Thus for an approximation to the oblate spheroid write

$$a^2 = \alpha^2(1+e), \quad b^2 = \alpha^2(1-e),$$

so that

$$(a^2+\lambda)(b^2+\lambda) = (\alpha^2+\lambda)^2 - e^2\alpha^4,$$

where $e^2$ is small. Then

$$F = 2\pi\rho abc\int_0^\infty \frac{d\lambda}{(\alpha^2+\lambda)^3(c^2+\lambda)^{\frac{3}{2}}}\left[1+\frac{3e^2\alpha^4}{2(\alpha^2+\lambda)^2}+\frac{15e^4\alpha^8}{8(\alpha^2+\lambda)^4}+\dots\right],$$

or with

$$\alpha^2+\lambda = (\alpha^2-c^2)(\nu'^2+1), \quad c^2+\lambda = (\alpha^2-c^2)\nu'^2, \quad \alpha^2 = (\alpha^2-c^2)(\nu^2+1), \quad c^2 = (\alpha^2-c^2)\nu^2,$$

$$F = \frac{4\pi\rho abc}{(\alpha^2-c^2)^{\frac{7}{2}}}\int_\nu^\infty \frac{d\nu'}{\nu'^2(\nu'^2+1)^3}\left[1+\frac{3e^2(\nu^2+1)^2}{2(\nu'^2+1)^2}+\frac{15e^4(\nu^2+1)^4}{8(\nu'^2+1)^4}+\dots\right]$$

$$= \frac{4\pi\rho\sqrt{1-e^2}}{(\alpha^2-c^2)^2}\,\nu(\nu^2+1)\int_\nu^\infty \frac{d\nu'}{\nu'^2(\nu'^2+1)^3}\left[1+\frac{3e^2(\nu^2+1)^2}{2(\nu'^2+1)^2}+\frac{15e^4(\nu^2+1)^4}{8(\nu'^2+1)^4}+\dots\right].$$

The integrations are effected by the use of

$$q_n = \nu(\nu^2+1)^{n-1}\int_\nu^\infty \frac{d\nu'}{\nu'^2(\nu'^2+1)^n},$$

in which

$$q_1 = 1-\nu\cot^{-1}\nu, \text{ and } (2n+1)(\nu^2+1)q_n - 2nq_{n+1} = 1 \quad \dots\dots\dots\dots(24)$$

is a sequence equation. Thus

$$\left.\begin{aligned} F &= \frac{4\pi\rho\sqrt{1-e^2}}{(\alpha^2-c^2)^2(\nu^2+1)}\left[q_3+\frac{3e^2}{2}q_5+\frac{15e^4}{8}q_7\dots\right] \\ G &= \frac{4\pi\rho\sqrt{1-e^2}}{\alpha^2-c^2}\left[q_2-q_3+\frac{3e^2}{2}(q_4-q_5)+\dots\right] \\ H &= 4\pi\rho\sqrt{1-e^2}(\nu^2+1)\left[q_1-2q_2+q_3+\frac{3e^2}{2}(q_3-2q_4+q_5)+\dots\right] \end{aligned}\right\} \quad \dots\dots\dots(25).$$

Since

$$a^2c^2+b^2c^2-a^2b^2 = (\alpha^2-c^2)^2(\nu^2+1)\{\nu^2-1+e^2(\nu^2+1)\},$$

and $c^2 = (\alpha^2 - c^2)\nu^2$, the Jacobian condition (9) written to the order $e^4$ is

$$0 = \nu^2 (q_1 - q_2) - (q_2 - q_3) + e^2 \left[\frac{3}{2}\{\nu^2 (q_3 - q_4) - (q_4 - q_5)\} + (\nu^2 + 1)(q_2 - q_3)\right]$$

$$+ e^4 \left[\frac{15}{8}\{\nu^2 (q_5 - q_6) - (q_6 - q_7)\} + \frac{3}{2}(\nu^2 + 1)(q_4 - q_5)\right] + \dots \quad \dots\dots(26\,a),$$

and
$$\frac{\omega^2}{4\pi\rho\sqrt{1 - e^2}} = \nu^2 (q_1 - q_2) + q_1 - 2q_2 + q_3 + \frac{3e^2}{2}\{\nu^2 (q_3 - q_4) + q_3 - 2q_4 + q_5\} + \dots$$

$$= q_1 - q_2 + e^2 \left\{\frac{3}{2}(q_3 - q_4) - (\nu^2 + 1)(q_2 - q_3)\right\} + e^4 \left\{\frac{15}{8}(q_5 - q_6) - \frac{3}{2}(\nu^2 + 1)(q_4 - q_5)\right\} + \dots \quad \dots\dots\dots(26\,b),$$

on using the equation of condition.

The term in (26 $a$) not containing $e$ vanishes for $\nu = \nu_0$ the value at the initial point; $e^2$ and the fall in $\omega^2$ are of the order $\nu_0 - \nu$. The condition of equal volumes makes

$$a_0^4 c_0^2 = a^2 b^2 c^2 = \alpha^4 c^2 (1 - e^2),$$

while $\alpha^2 : c^2 = \nu^2 + 1 : \nu^2$, so that $\alpha^2$ and $c^2$ differ from $a_0^2$ and $c_0^2$ also by quantities of order $\nu_0 - \nu$ or $e^2$, but $a$ and $b$ differ by a quantity of order $e$.

For a single calculation a value of $\psi$ in slight excess above 54° 21′ 27″·45 may be taken, then

$$q_1 = 1 - \psi \cot \psi = 1 - \nu \cot^{-1} \nu,$$

while $q_2 \dots$ are given by the sequence equation. A method applicable to all forms deviating only slightly from the initial is to find differential coefficients, and calculate their values for $\nu = \nu_0$. Differentiation is made by

$$\nu \frac{dq_n}{d\nu} = (2n - 1)(q_n - q_{n-1}), \quad \nu \frac{dq_1}{d\nu} = q_1 - \frac{1}{\nu^2 + 1} \quad \dots\dots\dots\dots\dots\dots\dots\dots(27),$$

and we may note that with the Jacobian condition

$$\frac{d}{d\nu}\{\nu^2 (q_1 - q_2) - (q_2 - q_3)\} = 2\nu q_1 (1 - \nu^2)/(1 + \nu^2), \quad \frac{d}{d\nu}(q_1 - q_2) = \frac{4}{\nu(\nu^2 + 1)}\{-1 + q_1 (4\nu^2 + 1)\}.$$

When evaluation is carried to the second order of $\Delta\psi$ and fourth order of $e$,

$$\left.\begin{aligned} e^2 &= 3{\cdot}445813\Delta\psi - 2{\cdot}78454(\Delta\psi)^2 \dots \\ \frac{\omega^2}{4\pi\rho} &= {\cdot}09355743 - {\cdot}1294005\Delta\psi - {\cdot}087816(\Delta\psi)^2 \dots \end{aligned}\right\} \quad \dots\dots\dots\dots\dots\dots(28).$$

Here $\Delta\psi$ is in circular measure, and values up to 30″ carry us approximately as far as the position $\alpha = 15°$ in Darwin's notation (*v. infra* p. 79).

§ 8. It may be of service to add expressions in terms of $q_1$ and $\nu^2$ for various quantities used in dealing with a spheroid $a, a, c$ where $a^2 : c^2 = \nu^2 + 1 : \nu^2$. Thus

$$F = \frac{\pi\rho}{2(a^2 - c^2)^2 (\nu^2 + 1)}\{15(\nu^2 + 1)^2 q_1 - 5\nu^2 - 7\},$$

$$G = \frac{\pi\rho}{2(a^2 - c^2)}\{5\nu^2 + 3 - 3(\nu^2 + 1)(5\nu^2 + 1) q_1\},$$

$$H = \frac{\pi\rho(\nu^2 + 1)}{2}\{(15\nu^4 + 6\nu^2 - 1) q_1 - 5\nu^2 + 1\},$$

which satisfy a relation

$$2Fa^2c^2 = G(2c^2 - a^2) + 3H \qquad \text{.........(29)}.$$

Also $\quad A = B = 2\pi\rho\{1-(\nu^2+1)q_1\}, \quad C = 4\pi\rho(\nu^2+1)q_1,$

$$\omega^2 = 2\pi\rho\{1-(3\nu^2+1)q_1\}, \quad \text{and} \quad -E = Aa^2 + Cc^2 = 2\pi\rho a^2\{1+(3\nu^2-1)q_1\}.$$

The equatorial oscillation of the spheroid has $\delta a + \delta b = 0$, and

$$n_e^2 = E_{aa} - E_{ab} = 4\pi\rho\{(3\nu^4+8\nu^2+1)q_1 - \nu^2 - 1\};$$

the polar oscillation has $\delta a = \delta b$, and

$$n_p^2(3\nu^2+1) = (\nu^2+1)(E_{aa}+E_{ab}) = 2\pi\rho(\nu^2+1)\{(27\nu^4+18\nu^2-1)q_1 - 9\nu^2+1\} \qquad \text{...(30)}.$$

The Jacobian condition in the form (7) or (9) leads to

$$\nu^2 + 3 = (3\nu^4 + 14\nu^2 + 3)q_1 \qquad \text{.........(31)},$$

and it may be verified that this makes $n_e^2 = 4\omega^2$. Or to establish this in a way which shews more clearly the points on which the exact relation turns, we may take the value of $E_{aa} - E_{ab}$ in the general form (10), which for $a = b$ is

$$\{2Fa^4c^2 + 4Ga^2c^2 + H(a^2+c^2)\}/a^2.$$

If we now combine the Jacobian condition at the junction

$$Ga^2(2c^2 - a^2) + Hc^2 = 0,$$

with the relation (29) true for all oblate spheroids, we have

$$2Fa^4c^2 = H(3a^2 - c^2).$$

This makes $\quad E_{aa} - E_{ab} = 4(Gc^2 + H) \quad \text{or} \quad n_e^2 = 4\omega^2.$

§ 9. The treatment at the other extremity depends on its approximation to the form of the prolate spheroid. For the latter with axes $a, c, c$ and $a^2 : c^2 = \nu^2 : \nu^2 - 1$,

and $$p_1 = \frac{\nu}{2}\log_e\frac{\nu+1}{\nu-1} - 1, \quad (2n+1)(\nu^2-1)p_n + 2np_{n+1} = 1 \qquad \text{.........(32)};$$

the formulae for $F, G, H$ are

$$\frac{4\pi\rho p_3}{c^2(a^2-c^2)}, \quad \frac{4\pi\rho(p_2-p_3)}{a^2-c^2}, \quad \frac{4\pi\rho(p_1-2p_2+p_3)c^2}{a^2-c^2}.$$

For the elongated Jacobian write

$$b^2 = \gamma^2(1+e), \quad c^2 = \gamma^2(1-e), \quad a^2 = (a^2-\gamma^2)\nu^2, \quad c^2 = (a^2-\gamma^2)(\nu^2-1),$$

then proceeding as for the oblate spheroid

$$F = \frac{4\pi\rho\sqrt{1-e^2}}{(a^2-\gamma^2)^2(\nu^2-1)}\left\{p_3 + \frac{3e^2}{2}p_5 + \frac{15e^4}{8}p_7 + \dots\right\}, \quad G \dots.$$

We are concerned with the case in which $\nu$ approaches the limit 1, say $\nu^2 = 1 + \xi^2$, and note that while $p_1$ becomes infinite $p_2, p_3 \dots$ approach limits $\frac{1}{2}, \frac{1}{4}, \frac{1}{6} \dots$ and approximately

$$F = \frac{4\pi\rho}{(a^2-\gamma^2)^2\xi^2}\left(p_3 + \frac{e^2}{16}\right), \quad G = \frac{4\pi\rho}{a^2-\gamma^2}\left(p_2 - p_3 - \frac{e^2}{16}\right), \quad H = 4\pi\rho\,\xi^2\left\{p_1 - 2p_2 + p_3 - \frac{e^2}{2}\left(p_1 - \frac{7}{8}\right)\right\}$$

$$\text{.........(33)}.$$

With $L = \log_e \frac{2}{\xi}$ we have as sufficient approximations

$$p_1 = L - 1 + \xi^2\left(\frac{L}{2}+\frac{1}{4}\right),\quad p_2 = \frac{1}{2} - \frac{3\xi^2}{2}(L-1),\quad p_3 = \frac{1}{4}\left(1-\frac{5\xi^2}{2}\right) \quad \text{..........(34)}.$$

The Jacobian condition $G(a^2c^2 + b^2c^2 - a^2b^2) + Hc^2 = 0$ is then

$$\left(p_2 - p_3 - \frac{e^2}{16}\cdots\right)(-2e + \xi^2 - 2e\xi^2\ldots) + \left\{p_1 - 2p_2 + p_3 - \frac{e^2}{2}\left(p_1 - \frac{7}{8}\right)\right\}\xi^2(1-e) = 0,$$

the solution of which gives

$$e = \xi^2(2L-3) + \xi^4\left(8L^2 - 20L + \frac{31}{2}\right) + \ldots \quad \text{..................(35 a)},$$

and then

$$\omega^2 = Gc^2 + H = 2\pi\rho\xi^2\{2L - 3 + \xi^2(3L-1) + \ldots\} \quad \text{................(35 b)}.$$

The first approximation uses only the main terms of $G$, $H$, the second involves an advance of one power of $e$, so that only the main terms of $G$ and $H$ are here required. Values of $E_{aa}$, ... sufficient for our purpose are:

$$a^2E_{aa} = 4\pi\rho(a^2-\gamma^2)\xi^2\left[L + \frac{1}{4} - \xi^2\left(\frac{9L}{2} - \frac{61}{8}\right)\cdots\right],$$

$$b^2E_{bb} = 4\pi\rho(a^2-\gamma^2)\xi^2\left[1 - \frac{\xi^2}{2}\cdots\right],$$

$$abE_{ab} = 4\pi\rho(a^2-\gamma^2)\xi^2\left[\frac{1}{2} - \xi^2\left(2L - \frac{11}{4}\right)\cdots\right].$$

One root of the quadratic in $n^2$ to which (1) leads is then $n_e^2 = 4\pi\rho\xi^2L$, the other is

$$n_p^2 = 2\pi\rho\left(1 - \frac{\xi^2}{2}\cdots\right).$$

Thus $n_e^2 : \omega^2$ falls from the value 4 to 1 in passing through the whole Jacobian series, i.e. this oscillation changes from a semi-daily to a daily oscillation. For $n_e$ the ultimate ratios of amplitudes are $-\frac{\delta a}{2a} = \frac{\delta b}{b} = \frac{\delta c}{c}$; for $n_p$, $\frac{\delta b}{b} + \frac{\delta c}{c} = 0$ to the first order, and $\frac{\delta a}{a}$ is of higher order $= -\xi^4(2L-3)\frac{\delta b}{b}$ approximately. These ratios answer to the description given above.

The Jacobian algebraical condition $a^2b^2 > c^2(a^2+b^2)$ is expressed by $x^2+y^2<1$.

In the limit $x^2=\xi^2$ and $y = 1 - e = 1 - \xi^2(2L-3)$, from which $x^2+y^2 = 1 - \xi^2(4L-7)$, i.e. the circle is approached: an approximate equation to the curve near $x=0$, $y=1$ is $1 - y = x^2\left(2\log_e\frac{2}{x} - 3\right)$. The positions for $\xi = \cdot 05$, and $\xi = \cdot 02$ are entered in the table (p. 80).

Darwin pointed out that there is a maximum value for kinetic energy which falls in this range. Assuming that it does lie here, the value of $(a^2+b^2)\,\omega^2$ is to be made a maximum as dependent on the single variable $\xi$. With $abc = 1$, $a^2+b^2 = \xi^{-\frac{4}{3}}\left(1 + \frac{5\xi^2}{3}\right)$ and we have to make $\xi^{\frac{2}{3}}\left[2L - 3 + \xi^2\left(\frac{19L}{3} - 6\right)\right]$ a maximum, the condition being

$$L - 3 + \xi^2\left(\frac{38L}{3} - \frac{67}{4}\right) = 0.$$

As $L=3$ roughly $\xi$ is not very small, on the margin of applicability of the formula. But the correction given by the second term is a small fraction of the prime value, and $\xi = \cdot 13791$ giving $e = \cdot 051624$ is probably near the mark: it corresponds well with the place indicated by Darwin's table ($D$ in fig. 1).

§ 10. We proceed to examine the boundaries of the domain of steady motion about one axis. The condition $\tau = 0$ is expressed by

$$(Aa^2 - Cc^2)/a = (Bb^2 - Cc^2)/b, \quad \text{or} \quad G(abc^2 + a^2c^2 + b^2c^2 - a^2b^2) + H(ab + c^2) = 0 \ldots (36).$$

The algebraical limit required to make the coefficient of $G$ negative with $c/a = x$, $c/b = y$ is $x^2 + xy + y^2 < 1$; the form $(ab + c^2)(H - abG) + c^2(a+b)^2 G = 0$, shews that $abG > H$ for this line. The ellipse $x^2 + xy + y^2 = 1$ is in fact much nearer to the Jacobian line. The inequality $x^2 + xy + y^2 > 1$ which appears to hold for the latter is evidently true if $abG > H$, but I have not found any proof of either inequality apart from the calculations on which the diagram is based. We accept Riemann's calculation for the terminus of $\tau = 0$ on $OS$, and treatment similar to that for the Jacobian gives

$$e = \xi(2L - 3) + \xi^2(L - 1)(2L - 3) + \ldots \quad \ldots\ldots\ldots\ldots (37),$$

for the terminus at $P$. The first approximation to its equation near $P$ is

$$1 - y = x\left(2\log_e \frac{2}{x} - 3\right).$$

For intermediate positions the expression of (36) by elliptic integrals has been used, viz.

$$(F - E)\cot^2\alpha\cos\beta(1 + \cos\beta) - E\{\cos^2\beta + \cos^2\gamma(1 - \cos\beta)\} + \cos\beta\sin\gamma\cos\gamma(2 - \cos\beta) = 0 \quad \ldots\ldots (38).$$

At the boundary $\tau' = 0$ the mean axis is axis of rotation, and if we retain the order of magnitude $a > b > c$, the condition is

$$(Aa^2 - Bb^2)/a = (Bb^2 - Cc^2)/c \quad \text{or} \quad G(acb^2 - a^2b^2 - b^2c^2 + a^2c^2) + H(ac - b^2) = 0 \ldots (39).$$

The first form of the condition shews that $b$ the axis of rotation is intermediate between $a$ and $c$. If the axis of rotation is taken for numerator of the fractions used as coordinates in the diagram $X = b/a$, $Y = b/c$ the field for this case is above the line $SP$ where it adjoins the field for least axis as axis of rotation. In the form $(ac - b^2)(acG + H) - b^2(a - c)^2 G = 0$ it appears that $ac > b^2$ or $1 > XY$, and the original form then involves $b^2(a^2 + c^2) > ac(ac + b^2)$ or $X^2 - XY + Y^2 > 1$; the boundary therefore lies between $XY = 1$ and $X^2 - XY + Y^2 = 1$. The graph shews the inequality $X + Y < 2$, or $b$ less than the harmonic mean of $a$ and $c$.

Near the terminus at $P$, working as for the Jacobian, we find

$$e = \xi(2L - 3) - \xi^2(L - 1)(2L - 3) \ldots \quad \ldots\ldots\ldots\ldots (40),$$

and the equation to the curve near $P$ is $Y = 1 + X\left(2\log_e \frac{2}{X} - 3\right)$.

The expression of (39) by elliptic integrals

$$(F - E)\cot^2\alpha\{\cos\gamma + \cos^2\beta(1 + \cos\gamma)\}$$
$$- E\{\cos^2\gamma + \cos^2\beta(1 + \cos\gamma)\} + \cos\beta\sin\gamma\cos\gamma(2 + \cos\gamma) = 0 \ldots (41),$$

with $\alpha > 45^\circ$, has been used for intermediate positions.

Near the spherical terminus of the boundary $\tau'=0$ we may write $a^2=1+\alpha', \ldots$ and then $P=(1+\lambda)^3+(1+\lambda)^2\Sigma\alpha'+(1+\lambda)\Sigma\alpha'\beta'+\alpha'\beta'\gamma'$. Proceeding to the second order of small quantities and using $\Sigma\alpha'+\Sigma\alpha'\beta'=0$, $P=(1+\lambda)^3-\lambda(1+\lambda)\Sigma\alpha'\beta'=\mu^3-\mu(\mu-1)\Sigma\alpha'\beta'$ where $\mu=1+\lambda$. Thus $A=2\pi\rho\int_0^\infty \frac{d\lambda}{(a^2+\lambda)\sqrt{P}}=2\pi\rho\int_1^\infty \frac{d\mu}{\mu^{\frac{5}{2}}}\left\{1-\frac{\alpha'}{\mu}+\frac{\alpha'^2}{\mu^2}+\left(\frac{1}{2\mu}-\frac{1}{2\mu^2}\right)\Sigma\alpha'\beta'\ldots\right\}$ or

$$A/4\pi\rho=\frac{1}{3}-\frac{\alpha'}{5}+\frac{\alpha'^2}{7}+\frac{\Sigma\alpha'\beta'}{35}\ldots \qquad (42).$$

The condition (39) then gives $\gamma'+\alpha'=\frac{9\alpha'^2}{7}$, $\beta'=-\frac{2\alpha'^2}{7}$; the first approximation to the curve near $S$ is $X+Y=2$, the second is $3(X^2+Y^2)+XY=7$, and the initial value of $\alpha$ for use in the elliptic functions is 45°.

§ 11. If we wish to shew what shapes of ellipsoid admit of rotation about either axis, we transfer to coordinates $x$, $y$ having $x=c/a$, $y=c/b$ so that $Yy=1$, $x=X/Y$, $X=x/y$.

The curve $\tau'=0$, $SHP$ in fig. 2, starts from $S$ along the line $x+1=2y$ which corresponds to $X+Y=2$. Near $P$ its equation is $1-y=x(2\log_e 2/x-3)$ corresponding to

$$Y=1+X(2\log_e 2/X-3);$$

i.e. the transformed curve approaches $P$ touching the line $\tau=0$ but above it as shewn by terms of the next order. On this transformed boundary $\tau'$ is not zero for the motion in which the *least* axis is that of rotation. The region $SHP$ represents ellipsoids capable of either motion, but with different values of $\tau$ and $\tau'$ except along the line $SP$: in particular the Jacobian ellipsoids of the range $HP$ are all capable of steady motion about the mean axis.

At $H$ we have $Aa^2-Bb^2=(a^2-b^2)(Bb^2-Cc^2)/b^2$ as Jacobian condition, while for $\tau'=0$ in the alternative motion $Aa^2-Bb^2=a(Bb^2-Cc^2)/c$, and therefore

$$c(a^2-b^2)=b^2a \quad \text{or} \quad b^2=a^2c/(a+c),$$

i.e. $y^2=x(1+x)$ or $\cos\gamma\tan^2\beta=1$. [Working through the $G$ and $H$ relation we get also a factor $b^2-c^2$ pointing to intersection at $P$.] The position of $H$ is near $x=\cdot4026$, $y=\cdot7515$, or $\alpha=67°\,17'$, $\gamma=66°\,15'\,24''$, least axis $\frac{2}{5}$ of the greatest. The $\tau$ of the one motion and $\omega$ of the other are connected by $\tau^2=\omega^2a^2(a-c)^3/2(a+c)$, the kinetic energies are in the ratio $a-c:a+2c$, the squares of momenta as $(a-c)^3(a+c):2a^2(a+2c)^2$. Numerically the kinetic energy of the motion about mean axis is 33·1 % of that for the Jacobian, the angular momentum $21\frac{2}{3}$ %; so that the disparity is considerable.

In the region of rotation about mean axis lies the case of irrotational motion to which Sir Alfred Greenhill called attention. The condition is

$$(Aa^2-Bb^2)/a^2(a^2+3c^2)=(Bb^2-Cc^2)/c^2(3a^2+c^2)$$

or

$$c^2(a^2-b^2)(3a^2+c^2)(Gc^2+H)=a^2(b^2-c^2)(a^2+3c^2)(Ga^2+H).$$

With respect to this I find that near $S$ it agrees with $\tau'=0$ to the second order of small quantities. Near $P$ I get

$$2e[p_2-p_3+\xi^2\{5(p_2-p_3)+p_1-p_2\}=\xi^2[3(p_1-p_2)-3e(p_1-p_3)+4\xi^2(p_1-p_2)],$$

leading to

$$e=3\xi^2(2L-3)-\xi^4\left(24L^2-56L+\frac{39}{2}\right)\ldots.$$

This gives a contact with $SP$ at $P$ similar to that of the Jacobian, below $SP$ in the standard system, above it in the other, and the sign of the second term suggests the point of inflexion necessary to secure this tangency.

§ 12. The table of Jacobians is calculated for 5° intervals of the variable $\alpha$, those for $\alpha = 5°$, and $\alpha = 10°$ by a method which will shortly be given, the rest by elliptic integrals; beyond these are added values for the place of maximum kinetic energy, also for $\xi = \cdot 05$ and $\xi = \cdot 02$ in the extreme range. The object in choosing $\alpha$ for the regular intervals rather than $\gamma$, as in Darwin's table, is to secure a safe interpolation. Gauss's form may with advantage be applied to tabulated integrals in which the integrand is calculable or tabulated more fully than the integral. Where Taylor's theorem holds good within a short interval $a$ to $a+b$,

$$\left.\begin{aligned}\int_a^{a+b} u_x dx &= b\left[u_a + \frac{b}{2}u_a' + \frac{b^2}{6}u_a'' + \frac{b^3}{24}u_a''' + \dots\right]\\ &= bu_{a+\frac{1}{2}b} + \frac{b^3 u_a''}{24} + \dots\\ \text{or}\qquad &= \frac{b}{2}(u_{a+\lambda b} + u_{a+\lambda' b}) + \frac{b^5 u_a^{\text{IV}}}{4320} + \dots\end{aligned}\right\}\dots\dots(43),$$

where $\lambda + \lambda' = 1$, $\lambda^2 + \lambda'^2 = \frac{2}{3}$, i.e. $\lambda = \cdot 211325\dots$, $\lambda' = \cdot 788675\dots$. The form $bu_{a+\frac{1}{2}b}$ is sufficient for a short range, $\frac{b}{2}(u_{a+\lambda b} + u_{a+\lambda' b})$ covers a much wider range. Thus requiring an elliptic integral with 60° 28′ for $\gamma$ we should write $\gamma' = 60°\ 14'$, use $\sin\beta' = \sin\alpha\sin\gamma'$, and then with $\Delta\gamma$ for the circular measure of 28′ the simpler corrections are $\Delta\gamma\cos\beta'$ for $E$ and $\Delta\gamma/\cos\beta'$ for $F$ to be added to the tabulated values for $\gamma = 60°$.

In the columns giving $(a^2+b^2)\omega^2$, $(a^2+b^2)^2\omega^2$, and $E$, use is made of $abc = 1$; the transfer to any other scale of magnitude is well understood.

On completing the work I was anxious to discover the source of the discrepancy between Darwin's results and my own as to the maximum velocity, a discrepancy occurring where the use of elliptic integral tables is troublesome and uncertain, viz. for smaller values of $\alpha$; and this paragraph is a sketch of work undertaken with that object. The forms in (15) and (16) are expanded in powers of $\sin^2\alpha$, the method is therefore closely allied to, though not identical with, Darwin's treatment given in a long footnote. The expansion of the left-hand member of (15) has the form $C_4\sin^4\alpha + C_6\sin^6\alpha + \dots$, where

$$C_4 = \sin^4\gamma\tan\gamma\,\{q_1(3\nu^4 + 14\nu^2 + 3) - \nu^2 - 3\}/8,$$

with $\nu = \cot\gamma$ and $q_1 = 1 - \gamma\cot\gamma$. The bracket vanishes at the junction by (31); near it the value of $C_4$ is $\frac{dC_4}{d\gamma}\Delta\gamma$, and a first approximation gives $\frac{dC_4}{d\gamma}\Delta\gamma + C_6\sin^2\alpha = 0$. We find

$$\frac{dC_4}{d\gamma} = 2q_1\sin^2\gamma(\cos^2\gamma - \sin^2\gamma),$$

and

$$C_6 = \sin^6\gamma\tan\gamma\,\{q_1(3\nu^6 + \nu^4 + 33\nu^2 + 3) - (\nu^2+1)(\nu^2+3)\}/32.$$

which at the junction is reduced to

$$C_6 = -q_1 \sin^3\gamma \cos\gamma (\cos^2\gamma - \sin^2\gamma)/2.$$

Thus the first approximation is $\Delta\gamma = \frac{1}{4}\sin^2\alpha \sin\gamma\cos\gamma$, which admits of immediate interpretation. For $x = \cos(\gamma + \Delta\gamma) = \cos\gamma\left(1 - \frac{1}{4}\sin^2\alpha\sin^2\gamma\right)$, and

$$y = \frac{\cos(\gamma+\Delta\gamma)}{\cos\beta} = \cos\gamma\left(1 + \frac{1}{4}\sin^2\alpha\sin^2\gamma\right),$$

and so $x + y = 2\cos\gamma = 2x_0$, i.e. the first approximation to the line near the junction shews a direction perpendicular to $OS$.

If the right-hand member of (16) is expanded as $D_0 + D_2 \sin^2\alpha + \ldots$ then

$$2D_0\cos\gamma = 1 - q_1(3\nu^2+1), \quad 16D_2\cos\gamma = \sin^2\gamma\,\{5\nu^2 + 3 - 3q_1(\nu^2+1)(5\nu^2+1)\},$$

or when simplified by the condition at the junction

$$4D_2\cos\gamma = \sin^2\alpha\cos^2\gamma\,\{1 - q_1(3\nu^2+1)\}.$$

Thus as far as $\sin^2\alpha$

$$\Delta\frac{\omega^2}{4\pi\rho} = \Delta(D_0\cos\gamma) + \frac{\sin^2\alpha}{4}(\cos^2\gamma - \sin^2\gamma)\,\{1 - q_1(3\nu^2+1)\},$$

and the first term is $\Delta\gamma\frac{d}{d\gamma}(D_0\cos\gamma)$ with the value of $\Delta\gamma$ just found. We have

$$\frac{d}{d\gamma}D_0\cos\gamma = \frac{\nu}{6}[q_1(27\nu^4 + 30\nu^2 + 3) - 9\nu^2 - 3],$$

which at the junction $= 4\nu\,\{q_1(3\nu^2+2) - 1\}$; and therefore

$$\Delta\frac{\omega^2}{4\pi\rho} = \sin^2\alpha\sin^2\gamma\,\{q_1(3\nu^4 + 14\nu^2 + 3) - \nu^2 - 3\}/12,$$

which vanishes at the junction.

Thus the first term in $\Delta\omega^2$ depends on $\sin^4\alpha$. The details of this method are very tedious, and I have not applied it to obtain the coefficient of $\sin^4\alpha$. The position is that $\omega^2$ with respect to the variable $\alpha$ is a sustained maximum, with respect to $\gamma$ an ordinary maximum, the variation depending on $\sin^4\alpha$ or on $(\Delta\gamma)^2$. [$\Delta\psi$ is of order $e^2$ or $\sin^4\alpha$ or $(\Delta\gamma)^2$.]

Connecting this with the method of approximation in § 7, with $\gamma$ for the value at the junction

$$e = \frac{1}{2}\sin^2\alpha\sin^2\gamma\left(1 + \frac{1}{2}\sin^2\alpha + \frac{1}{4}\sin^4\alpha\right) + \frac{\sin^6\alpha\sin^5\gamma\cos\gamma}{4\times 3{\cdot}445813} + \ldots,$$

$$\Delta\gamma = \frac{1}{4}\sin^2\alpha\sin\gamma\cos\gamma + \frac{\sin^4\alpha\sin\gamma\cos\gamma}{32}(3\cos^2\gamma + \sin^2\gamma)$$
$$+ \frac{\sin^6\alpha\sin\gamma\cos\gamma}{384}(15\cos^4\gamma + 10\sin^2\gamma\cos^2\gamma + 3\sin^4\gamma)$$
$$+ \frac{\sin^4\alpha\sin^4\gamma}{4\times 3{\cdot}445813}\left\{1 + \frac{\sin^2\alpha}{4}(5\cos^2\gamma + 3\sin^2\gamma)\right\} + \ldots$$
$$= {\cdot}1183906\sin^2\alpha + {\cdot}0564942\sin^4\alpha + {\cdot}03562\sin^6\alpha + \ldots,$$

or in minutes of arc $= 407{\cdot}00\sin^2\alpha + 194{\cdot}21\sin^4\alpha + 122{\cdot}45\sin^6\alpha + \ldots$

and $$\frac{\omega^2}{4\pi\rho} = {\cdot}0935574 - {\cdot}0040948(\sin^4\alpha + \sin^6\alpha) - {\cdot}0009615\sin^8\alpha \ldots \qquad (44).$$

The last pair give a solution in terms of the angle $\alpha$ for all small values, more trustworthy I think than results derived from the direct solution* of (15) by tables, up to $\alpha = 15^\circ$.

*Schedule of solutions for the Jacobian form of equilibrium.*

| $\alpha$ | $\gamma$ | $x$ | $y$ | $\omega^2/4\pi\rho$ | $(a^2-b^2)^2\omega^2/4\pi\rho$ | $(a^2-b^2)\omega^2/4\pi\rho$ | $-2E/4\pi\rho$ | $n_e^2/\omega^2$ | $n_p^2/\omega^2$ |
|---|---|---|---|---|---|---|---|---|---|
| 0° | 54° 21′ 27″·45 | ·582724 | ·582724 | ·0935574 | ·768872 | ·268203 | 1·681975 | 4 | 4·1182 |
| 5° | 54° 24′ 33″·6 | ·58199 | ·58346 | ·0935572 | ·768876 | ·268205 | | | |
| 10° | 54° 33′ 54″·6 | ·57978 | ·58567 | ·0935536 | ·768933 | ·268210 | | | |
| 15° | 54° 49′ 40″ | ·5760 | ·5894 | ·09354 | ·7692 | ·2682 | 1·6818 | 3·9873 | 4·1316 |
| 20° | 55° 11′ 40″ | ·5707 | ·5947 | ·09349 | ·7699 | ·2683 | | | |
| 25° | 55° 41′ | ·5638 | ·6016 | ·09339 | ·7714 | ·2684 | | | |
| 30° | 56° 17′ 40″ | ·5549 | ·6102 | ·09322 | ·7745 | ·2687 | 1·6800 | 3·8887 | 4·2447 |
| 35° | 57° 2′ 12″ | ·5441 | ·6207 | ·09291 | ·7797 | ·2692 | | | |
| 40° | 57° 55′ 51″ | ·5309 | ·6331 | ·09243 | ·7887 | ·2700 | | | |
| 45° | 58° 58′ | ·5156 | ·6481 | ·09162 | ·8017 | ·2710 | 1·6707 | 3·6876 | 4·5190 |
| 50° | 60° 11′ 30″ | ·4971 | ·6654 | ·09043 | ·8225 | ·2727 | | | |
| 55° | 61° 37′ 30″ | ·4752 | ·6856 | ·08868 | ·8543 | ·2753 | | | |
| 60° | 63° 18′ 5″ | ·4493 | ·7092 | ·08610 | ·9031 | ·2789 | 1·5385 | 3·1626 | 5·3315 |
| 65° | 65° 16′ 48″ | ·4182 | ·7367 | ·08238 | ·9798 | ·2841 | | | |
| 70° | 67° 38′ 30″ | ·3804 | ·7689 | ·07694 | 1·1054 | ·2916 | | | |
| 75° | 70° 29′ 41″ | ·3339 | ·8074 | ·06896 | 1·3251 | ·3023 | 1·2444 | 2·894 | 6·821 |
| 80° | 74° 5′ | ·2742 | ·8540 | ·05694 | 1·7685 | ·3173 | | | |
| 85° | 78° 59′ 6″ | ·1911 | ·9127 | ·03773 | 3·0055 | ·3368 | | | |
| | | ·1330 | ·9496 | ·02360 | 4·9407 | ·3412 (max.) | | | |
| | | ·0497 | ·9887 | ·00220 | 16·36 | ·3000 | | | |
| | | ·0200 | ·9975 | ·00124 | 42·23 | ·2291 | | | |
| 90° | 90° | 0 | 1 | 0 | $\infty$ | 0 | 0 | 1 | $\infty$ |

§ 13. We now propose to discuss the delimitation of boundaries for steady motion in which two axes are involved. Riemann curiously enough omitted to state the order of magnitude of the three axes for his several cases, and dealt only slightly with boundaries of transcendental type. In other respects his statement of the analytical conditions is so clear and precise that the most judicious course appears to be, to state his results in his own notation, then transfer, when the order is established, to a standard order $a > b > c$ and add what is necessary to complete the delimitation. Riemann uses $b$ and $c$ for axes of rotation in all cases.

CASE I. Greatest and least axes those of rotation. Riemann gives $b \geqslant a \geqslant \frac{b+c}{2}$. Here $\frac{b}{a} + \frac{c}{a} < 2$, and $\frac{b}{a} \geqslant 1$, therefore $\frac{c}{a} < 1$; thus the order is $b > a > c$ with a possible equality in the limit. For standard order we interchange $a$ and $b$ when the condition is $2b \geqslant a + c$, $a$ and $c$ being axes of rotation. The boundary $2b = a + c$ is represented by the hyperbola

* The difficulty in working with the tables at this point is that a small alteration in $\alpha$ or $\gamma$ gives a much less alteration in the residue of (15) than in (16) or in the individual terms of either. Thus a small error in the adjustment of $\alpha$ and $\gamma$ involves a larger error in $\omega^2$.

$y(1+x)=2x$, the part relevant is the arc from $O$ to $S$, and the vertex $V$ of the hyperbola $x=\sqrt{2}-1$, $y=2-\sqrt{2}$ lies symmetrically with regard to the chord $OS$. The region concerned is a band skirting Maclaurin's line, and the state of motion is continuous with Maclaurin's for $a=b$.

CASE II. Axes of rotation, greatest and least. Riemann takes $b>c$ and then

$$b-c>2a, \quad c^2<a^2(b^2-4a^2)/(b^2-a^2)$$

are conditions. As $b>2a$, $b^2-4a^2$ and $b^2-a^2$ are both positive and $c<a$, i.e. the order is $b>a>c$. The transfer to standard order as in Case I gives $a-c>2b$ and $c^2(a^2-b^2)<b^2(a^2-4b^2)$ as conditions, or $y(1-x)>2x$, $y^2-4x^2>y^2(y^2-x^2)$. The part of the quartic which is relevant is half of a loop which sets out from $O$ along the line $y=2x$, touches $SP$ at $P$, and has at $Q$ a maximum value for $x$ of value $2-\sqrt{3}$ where $y=\sqrt{3}-1$. The region is to the left of these lines, viz. within $OQP$; near $O$ the hyperbola imposes the more stringent condition, and this is the boundary till the point of intersection is reached at $Q$ where $x$ is a maximum. Beyond this point the quartic imposes the more stringent condition and is the boundary to $P$.

But this proves to be merely preliminary. The loop boundary is

$$0=4b^4-b^2(a^2+c^2)+a^2c^2\equiv D,$$

and $D$ is a factor of the determinant of Riemann's equations for $\sigma$, $S$, $T$. These equations have right-hand members, and with a zero determinant are not unconditionally consistent. The condition is $2Gb^2+H=0$, which is an impossible relation. The loop section of boundary is certainly not valid as a boundary *up* to which steady motion is possible. The intervention of a pressure condition, which in fact supersedes both algebraical conditions as the effective boundary, will be discussed below.

CASE III. Axes of rotation, the greatest and the mean. Riemann takes $b>c$, and then

$$2a\leqslant(b-c), \quad \int_0^\infty \frac{\lambda\,d\lambda}{P^{\frac{1}{2}}(b^2+\lambda)}\left\{\frac{4a^2-c^2+b^2}{c^2+\lambda}-\frac{b^2}{a^2+\lambda}\right\}\leqslant 0,$$

are the two inequalities to be satisfied.

If $c\geqslant 2a$ the integral inequality holds, and will continue to hold for values of $c$ less than $2a$ but not so small as $a$ unless $b$ becomes infinite. There is certainly a range within which the order is $b>c>a$, which is converted to standard order by a cyclical change, which makes the conditions

$$2c\leqslant(a-b), \text{ and } \int_0^\infty \frac{\lambda\,d\lambda}{P^{\frac{3}{2}}}\{a^2(b^2+\lambda)-(4c^2-b^2+a^2)(c^2+\lambda)\}\geqslant 0 \quad \text{.........(45).}$$

The first boundary $2c=a-b$ is $y(1-2x)=x$; the inequality $y(1-2x)>x$ is observed to the left of the hyperbola up to the point $T$ at which the integral equality is satisfied, for which we know that $1>y>\frac{1}{2}$. If this equality holds for a line $TP$ proceeding to $P$, the order remaining $a>b>c$, this line will complete the boundary. The equality

$$(a^2b^2+b^2c^2-a^2c^2-4c^4)\,G-(4c^2-b^2)\,H=0$$

meets the curve $\tau'=0$ or

$$(acb^2-a^2b^2-b^2c^2+a^2c^2)\,G+H\,(ac-b^2)=0,$$

for values making the eliminant

$$ac(b^2-c^2)(a-b-2c)(a+b-2c)$$

vanish, i.e. for $b^2=c^2$ giving the point $P$, and for $a=b+2c$ which is precisely the point $T$ where the two boundaries meet. Near $P$ we get

$$2e(p_2-p_3)+5e\xi^2(p_1-p_3)=3\xi^2(p_1-p_2),$$

and approximately $$e=3\xi^2(2L-3)-3\xi^4\left(8L^2-24L+\frac{29}{2}\right)\ldots.$$

Near $P$ the curve lies below the Jacobian, the type of first term is the same, the sign of the second approximation opposite. The Jacobian is met at the point $P$, and again for

$$a^2(3c^2-b^2)=c^2(4c^2-b^2) \quad \text{or} \quad x^2(4y^2-1)=3y^2-1$$

giving an approximate position $y=\cdot622$, $x=\cdot5417$; and the curve $PT$ must have a point of inflexion in order to cross the Jacobian. Solutions were obtained of the equation reduced to elliptic integrals, viz.

$$\sin^2\alpha\cos\beta\sin\gamma\cos\gamma(\sin^2\alpha+\sin^2\beta+4\cos^2\gamma)+2E\sin^2\alpha\cos^2\gamma(2-\sin^2\alpha)$$
$$=(E-F\cos^2\alpha)[\cos^2\gamma(4-\sin^2\alpha)+\cos^2\beta(\sin^2\beta+4\cos^2\gamma)]\ldots(46).$$

Along the whole length of $PT$ the transcendental boundary this case is continuous with the case of rotation about mean axis only. Here Riemann's $S=0$ makes $u$ and $u'$ vanish, and gives also $4T(4c^2-b^2)=G(b^2-c^2)$; the relation between $G$ and $H$ gives a ratio for $v^2:v'^2$ in the theory of mean axis which is clear of transcendentals, viz.

$$=(a+2c)^2-b^2:(a-2c)^2-b^2,$$

and the values of $v$, $v'$ for the two cases agree.

§ 14. The condition of positive pressure (or $\sigma$ positive) is one that is little in evidence, for the reason that in all cases of rotation about one axis as well as in Cases I and III the condition is satisfied as a simple consequence of other conditions. In Case II it restricts the field within narrower limits than the composite algebraical boundary, and the equation $p$ or $\sigma=0$ supersedes other conditions as the effective boundary of the case.

The pressure is positive so long as $F+3(2Gb^2+H)/D>0$, $D$ meaning

$$4b^4-b^2(a^2+c^2)+a^2c^2$$

with standard order. The condition $D=0$ is represented by the quartic loop, $D$ being negative between $OP$ and the loop. The line $p=0$ may be expected generally to lie well within the loop, but may meet it at $P$ or $O$ where other quantities than $D$ are small. Near $P$ the equality is represented by

$$p_3(2e-3\xi^2-6e\xi^2)=3\xi^2[2(p_2-p_3)(1+e)+p_1-2p_2+p_3]$$

which treated as before gives

$$e=6\xi^2(L-1)+9\xi^4\left(6L-\frac{11}{2}\right)+\ldots.$$

The curve near $P$ lies below the Jacobian or the boundary $PT$ of Case III, and turns more sharply downwards.

We examine now the position near $O$ where $c$ is small. The reciprocals of (5) give

$$F = \frac{A}{(a^2-b^2)(a^2-c^2)} - \frac{B}{(a^2-b^2)(b^2-c^2)} - \frac{C}{(a^2-c^2)(b^2-c^2)} \qquad (47)$$

while $-G$ and $H$ are got by writing $Aa^2$, $Aa^4$ respectively for $A$, .... Thus if $c$ is small the condition is reduced to

$$\frac{C}{a^2} - \frac{3\{A(a^2-2b^2)-Bb^2\}}{(a^2-b^2)(a^2-4b^2)} > 0.$$

Comparing the right-hand members of (14) $C$ has a term with $\cos\gamma$ in the denominator and is relatively great when $c$ is small, so that the positive character is assured unless $a^2-4b^2$ is small. But near the boundary $a-c=2b$ leads to $a^2-4b^2=2ac$, and in fact makes the above negative, but we can adjust $f$ in $a-fc=2b$ so as to satisfy the condition which is then

$$C - \frac{3\{A(a^2-2b^2)-Bb^2\}}{2f(a^2-b^2)\cos\gamma} > 0,$$

or as $a=2b$ approximately, the condition is $C > \frac{2A-B}{2f\cos\gamma}$. Now with $\gamma$ nearly $90°$ and $\beta=60°$ and therefore $\alpha=60°$, $C$ is proportional to $\frac{\sin^2\alpha\cos\alpha}{\cos\gamma}$ or $\frac{3}{8\cos\gamma}$, and $2A-B$ to $\frac{1}{4}(F_1+2E_1)$ where $E_1$ and $F_1$ are complete elliptic integrals for $\alpha=60°$. Hence the condition is

$$f > \tfrac{1}{3}(F_1+2E_1) > 1{\cdot}52619,$$

using $F_1=2{\cdot}15652$, $E_1=1{\cdot}21106$. The pressure line near $O$ is therefore $a-fc=2b$ or

$$y-2x=fxy=2fx^2=3{\cdot}0524x^2$$

approximately, whereas the hyperbolic boundary is $y=2x+2x^2$ near $O$; the curve then lies above the hyperbola but has the same tangent $y=2x$ at $O$, with a radius of curvature reduced as $2:3$.

For a general position we use

$$F(a^2-b^2)(b^2-c^2)-3b^4-3(2Gb^2+H)>0,$$

or the boundary is

$$Fa^4(\sin^2\beta\sin^2\gamma\cos^2\alpha-3\cos^4\beta)-3(2Ga^2\cos^2\beta+H)=0.$$

From (47) it follows that $Fa^4$, $-Ga^2$, $H$ are respectively proportional to

$$A\cos^2\alpha-B+C\sin^2\alpha,\quad A\cos^2\alpha-B\cos^2\beta+C\sin^2\alpha\cos^2\gamma,\quad A\cos^2\alpha-B\cos^4\beta+C\sin^2\alpha\cos^4\gamma.$$

In this way three points were determined in a middle range by use of elliptic integrals.

§ 15. Attention may be called briefly to some features of the steady motion about two axes, say $b$ and $c$ in Riemann's notation. The angular velocities $\omega_2$, $\omega_3$ and momenta $h_2$, $h_3$ are constant, and moreover $h_2 : h_3 = \omega_2 : \omega_3$; in fact we find

$$h_3 = \frac{m\omega_3}{10}\left[b^2+c^2-2a^2 \pm \sqrt{(2a+b+c)(2a-b-c)(2a+b-c)(2a-b+c)}\right],$$

the positive value of the radical applicable to the case in which $w$, $w'$ have like signs.

The frame of the ellipsoid moves in such a way that any line belonging to it describes a cone with uniform angular velocity about a line fixed in space. The axes in most parts of the fields for I, II and III are very unequal; the irrotational element in the motion must therefore be very influential in order to equalize the effective moments of inertia. In general the kinetic energy is not expressible in terms of $\omega$'s only, for $h_3$ being

$$\frac{m}{5}\{(a-b)^2 w + (a+b)^2 w'\}$$

and $\omega_3$ being $w + w'$, the kinetic energy

$$= \frac{m}{5}\{(a-b)^2 w^2 + (a+b)^2 w'^2 + (a-c)^2 v^2 + (a+c)^2 v'^2\}$$

$$= \frac{m}{5}[\{(a-b)^2 w + (a+b)^2 w'\}(w+w') - 2(a^2+b^2) ww' \ldots]$$

$$= h_2\omega_2 + h_3\omega_3 - \frac{2m}{5}\{(a^2+c^2) vv' + (a^2+b^2) ww'\}.$$

But when $w$, $w'$ have like signs, $v$ and $v'$ have unlike signs and the bracket *may* vanish, in which case the kinetic energy would be double that belonging to a rigid body with the momenta and velocities. In Riemann's notation

$$(a^2+c^2) vv' + (a^2+b^2) ww' = \{(a^2+b^2) T \pm (a^2+c^2) S\}$$
$$\times \sqrt{(2a+b+c)(2a-b-c)(2a+b-c)(2a-b+c)},$$

the positive sign attaching to Cases II and III where $S$ is negative, the negative sign to Case I where $S$ is positive. The condition $(a^2+b^2) T + (a^2+c^2) S = 0$ is represented by

$$6a^4 (Ga^2 + H) + H(a^2-b^2)(a^2-c^2) = 0,$$

which cannot be satisfied for III. But for II where in standard order the relation is

$$6b^4 (Gb^2 + H) = H(a^2-b^2)(b^2-c^2),$$

it is satisfied for a line from $O$ to $P$ within the domain of II. The line leaves $O$ within the region of positive pressure (angle $\alpha$ slightly less than 74°), quits this region in the upper part, but remains below the quartic in approaching $P$. To the left of the line the kinetic energy is less than the amount specified above, to the right it is greater. The existence of the line is a curious feature, which rather emphasizes the characteristics of this motion, but the line does not appear to have any true dynamical significance.

A consequence of the influential part played by irrotational motion in Case II is that the status would be greatly altered by a very small amount of friction. The reduction of mechanical energy would involve a movement towards the pressure boundary where cohesion ceases and the conditions become disruptive. If this approach takes place in the upper part of the diagram where there is approximation to the spindle shape, separation into two or three less elongated bodies seems probable. If the approach takes place in the lower part of the diagram where the form is that of a thin elliptical disk, the formation of a globe and ring seems more probable. That is we postulate transverse lines of weakness for the first case, an annular line (or lines) for the second. Without professing any special confidence in the application of a homogeneous fluid theory to cosmogony, it seems permissible to set down the above suggestions as those which arise most naturally from this part of the subject.

§ 16. The figure of equilibrium discussed by E. Roche, where a distant attracting body $M$ on the line of the greatest axis moves in a circle with an orbital period equal to that of rotation of the liquid ellipsoid, is represented by the equations

$$a^2 (A - \omega^2 - 2\nu\omega^2) = b^2 (B - \omega^2 + \nu\omega^2) = c^2 (C + \nu\omega^2),$$

where $\nu = M/(M+m)$ may range from 0 to 1. For $\nu = 0$ the distant body has an inappreciable influence, and the course of solutions is represented by the curve $SJP$ representing spheroids and Jacobians. The other extreme $\nu = 1$ is the case which has been specially examined. The general course for intermediate values is easy to forecast, and some help is given by the treatment near the spherical terminus after the method of § 10. Thus

$$(1-x)/(1-y) = 3\nu + 1 = \sec^2 \alpha$$

near $S$, i.e. the tangent of the inclination of curves to $SP$ near $S$ varies from 1 to $\frac{1}{4}$, and the initial value of the angle $\alpha$ used in the elliptic functions varies from 0° to 60°. The curves lie above these tangents and have one point of inflexion after reaching the position for which $y$ is a minimum. The first approximation near $S$ also makes $\dfrac{\omega^2}{4\pi\rho} = \dfrac{4(1-x)}{15(3\nu+1)}$.

For $\nu = 1$, the equation of condition is

$$\left.\begin{aligned} &\sin^2\alpha\cos\beta\sin\gamma\cos\gamma\,(6+\cos^2\gamma) - 2E\sin^2\alpha\cos^2\gamma - (E - F\cos^2\alpha)\{\cos^2\gamma + \cos^2\beta\,(3+\cos^2\gamma)\} = 0 \\ \text{and}\quad &\frac{\omega^2\cos^2\alpha\sin^2\beta\sin\gamma\cos\gamma}{4\pi\rho\cos\beta} = (E - F\cos^2\alpha)\cos^2\beta + E\sin^2\alpha\cos^2\gamma - 2\sin^2\alpha\cos\beta\sin\gamma\cos\gamma \end{aligned}\right\} \quad \ldots\ldots\ldots(48).$$

The table gives values for four positions

| $\alpha$ | $\gamma$ | $x$ | $y$ | $\omega^2/4\pi\rho$ |
|---|---|---|---|---|
| 70° | 43° 22′½ | ·72688 | ·95155 | ·016205 |
| 75° | 53° 45′ | ·59131 | ·94298 | ·021162 |
| 80° | 63° 17′½ | ·44945 | ·94528 | ·022400 |
| 85° | 73° 13′ | ·28875 | ·96069 | ·017705 |

The maximum value of $\omega^2/4\pi\rho$ inferred by interpolation is ·02253, the minimum value of $y$ ·9424, the place of the former nearer to $P$ than the latter.

## VI. *On the Fifth Book of Euclid's Elements* (Third Paper).

By M. J. M. Hill, M.A., LL.D., Sc.D., F.R.S.,
Astor Professor of Mathematics in the University of London.

[*Received* 1 September, 1914.]

### I. Preliminary.

1. The Cambridge Philosophical Society published two papers by me "On the Fifth Book of Euclid's Elements" in their *Transactions*, vol. xvi. part iv. (1897) and vol. xix. part ii. (1902). These will be referred to in what follows as my first and second papers. I shall also have occasion to refer to my two editions of the *Fifth and Sixth Books of Euclid* published by the Cambridge University Press in 1900 and 1908, and my *Theory of Proportion* published by Constable and Co. (1914).

### II. Object of this Paper.

2. The special object of this paper is to study the Fifth Book of Euclid from the point of view of its relation to the principle afterwards known as the Axiom of Archimedes.

I purpose to set out the results which can be obtained

(*a*) by considering this principle in connection with the Fourth and Fifth Definitions of the Fifth Book;

(*b*) by considering how far this principle is *necessarily* involved in the proofs of properties of *Equal* Ratios given in the Fifth Book.

### III. The Axiom of Archimedes and the Fourth Definition of the Fifth Book.

3. Though the principle is now always known as the Axiom of Archimedes it is very clearly assumed in the Fourth Definition of the Fifth Book, which Sir T. L. Heath translates as follows:

> Magnitudes are said to have a ratio to one another which are capable when multiplied of exceeding one another.

Thus it is assumed that if $A$ has a ratio to $B$, or $B$ to $A$, then it is always possible to find integers $n$ and $p$, such that $nA > B$ and $pB > A$.

This plainly assumes the Axiom of Archimedes, and it reads like an anachronism to call the axiom by its usual name, but I shall conform to the usual practice throughout this paper.

## IV. THE AXIOM OF ARCHIMEDES AND THE FIFTH DEFINITION OF THE FIFTH BOOK.

4. Euclid did not consider the bearing of the Axiom of Archimedes on the conditions enumerated in the Fifth Definition of the Fifth Book, but important results can be obtained by doing this.

Sir T. L. Heath translates the Fifth Definition as follows:

> Magnitudes are said to be in the same ratio, the first to the second and the third to the fourth, if any equimultiples whatever be taken of the first and third, and any equimultiples whatever of the second and fourth, the former equimultiples alike exceed, are alike equal to, or alike fall short of, the latter equimultiples respectively taken in corresponding order.

Hence $$A : B :: C : D,$$

if when we take any equimultiples whatever $rA$, $rC$ of $A$ and $C$: and any equimultiples whatever $sB$, $sD$ of $B$ and $D$, then the only relations which are *simultaneously* possible are

$$\text{(I)} \quad rA > sB, \quad rC > sD:$$

or $$\text{(II)} \quad rA = sB, \quad rC = sD;$$

or $$\text{(III)} \quad rA < sB, \quad rC < sD.$$

To make the meaning quite clear it may be added that (I) means that if having chosen the integers $r$, $s$ we find that $rA > sB$, it is necessary that $rC > sD$; and *also* that if we find $rC > sD$ it is necessary that $rA > sB$.

So that to express the full meaning of (I) we may say that if $r$, $s$ are any integers whatever such that

$$rA > sB, \text{ then must } rC > sD \qquad (1).$$

But if $r$, $s$ are such that

$$rC > sD, \text{ then must } rA > sB \qquad (2).$$

In like manner the full meaning of (II) is expressed by saying that if $r$, $s$ are any integers whatever such that

$$rA = sB, \text{ then must } rC = sD \qquad (3),$$

but if $$rC = sD, \text{ then must } rA = sB \qquad (4).$$

Similarly the full meaning of (III) is expressed by saying that if $r$, $s$ are any integers whatever such that

$$rA < sB, \text{ then must } rC < sD \qquad (5),$$

but if $$rC < sD, \text{ then must } rA < sB \qquad (6).$$

A proof is given in Art. 36 of my *Euclid V. and VI.* (1st edition) of the simple theorem that if the conditions (1), (3) and (5) are satisfied then (2), (4) and (6) can be deduced from them by reasoning which is purely *logical* and does not involve any knowledge of the properties of the magnitudes dealt with, *and in particular does not assume the truth of Archimedes' Axiom.* Euclid does not give a proof of this simple theorem, but assumes it throughout the whole of his Fifth Book. He does in fact reason as if he had defined the proportion

$$A : B :: C : D$$

as existing, when the magnitudes concerned satisfy the conditions (1), (3) and (5). Simson gives the definition in that form, but it is *unsymmetrically* related to the first and third of the magnitudes concerned, and also to the second and fourth, whilst Euclid's form is *symmetrical.*

In like manner it can be shewn that if the conditions (2), (4) and (6) are satisfied then (1), (3) and (5) can be deduced from them by purely *logical* processes.

It was further shewn in Art. 37 of the book just mentioned that if the conditions (1), (2), (5) and (6) are satisfied then (3) and (4) can be deduced from them by purely *logical* processes.

Further, it can be shewn that if either the condition (3) or (4) hold for a single value of $r$ and a single value of $s$, then all the remaining conditions hold good (see my *Theory of Proportion,* Art. 77). This particular case is mentioned for the sake of completeness, but it is of no importance in the argument. It can only occur when the magnitudes $A$ and $B$ are commensurable, and $C$ and $D$ are commensurable.

No other and no further reduction of the six conditions (1)—(6) to a smaller number by purely *logical* processes is possible.

5. But if the truth of Archimedes' Axiom is assumed it is possible to reduce the six conditions to

(1) and (5),

or to (2) and (6),

or to (1) and (2),

or to (5) and (6).

Stolz, in his *Vorlesungen über allgemeine Arithmetik,* part I. p. 87 (1885), was, I believe, the first to shew that (3) is involved in (1) and (5), and then (2), (4) and (6) follow.

This covers the case in which it is shewn that (4) is involved in (2) and (6), and then (1), (3) and (5) will follow.

Lastly, it will be shewn that (1) is equivalent to (6), and (2) to (5).

6. The reduction of the six conditions to (1) and (2) was I believe first given in my Second Paper, Art. 68, by a proof in which I shewed that they involved (5) and then the remainder followed by Stolz's work. A similar proof is of course applicable to the reduction of the six conditions to (5) and (6).

I will however now give a demonstration, independent of Stolz's work, due to my friend Mr Rose-Innes, of the reduction of the six conditions to (5) and (6).

Let $A$, $B$, $C$, $D$ be four magnitudes such that whenever

$rA < sB$, then $rC < sD$ ....................................(5),

and whenever $rC < sD$, then $rA < sB$ ....................................(6),

it will be proved that the remaining four conditions also hold.

The relation between the magnitudes $rA$, $sB$ must be one of the following:

$$rA < sB \quad \ldots\ldots\ldots\ldots (\alpha),$$
$$rA = sB \quad \ldots\ldots\ldots\ldots (\beta),$$
$$rA > sB \quad \ldots\ldots\ldots\ldots (\gamma).$$

The relation between the multiples $rC$, $sD$ must be one of the following:

$$rC < sD \quad \ldots\ldots\ldots\ldots (a),$$
$$rC = sD \quad \ldots\ldots\ldots\ldots (b),$$
$$rC > sD \quad \ldots\ldots\ldots\ldots (c).$$

Each of the cases of the first set must be considered in conjunction with each of the cases of the second set, so that there are nine combinations to be considered altogether. We have to reject all those which are inconsistent with (5) and (6), which we suppose to hold.

The possibility of the combination of $(\alpha)$ with $(b)$, and that of $(\alpha)$ with $(c)$, are inconsistent with (5).

The possibility of the combination of $(a)$ with $(\beta)$, and that of $(a)$ with $(\gamma)$, are inconsistent with (6).

Let us consider next the combination of $(\beta)$ with $(c)$, i.e. suppose

$$rA = sB,$$
$$rC > sD.$$

Then $rC - sD$ is a magnitude of the same kind as $D$, and now *introducing Archimedes' Axiom* we can assert the existence of an integer $n$, such that

$$n(rC - sD) > D,$$
$$\therefore\ nrC > (ns+1)D,$$

but since

$$rA = sB,$$
$$\therefore\ nrA = nsB < (ns+1)B,$$

so that the integers $(nr)$ and $(ns+1)$ are such that

$$(nr)A < (ns+1)B \text{ but } (nr)C > (ns+1)D,$$

which is inconsistent with condition (5).

Considering next the combination $(\gamma)$ with $(b)$, i.e. $rA > sB$, $rC = sD$, it can be shewn in the same way by interchanging $A$ with $C$ and $B$ with $D$ in the preceding proof that this is inconsistent with condition (6).

It has therefore been proved that the combinations

$(\alpha)$ with $(b)$,

$(\alpha)$ with $(c)$,

$(a)$ with $(\beta)$,

$(a)$ with $(\gamma)$,

$(\beta)$ with $(c)$,

and $(\gamma)$ with $(b)$

are inconsistent with the conditions (5) and (6).

Hence, if the conditions (5) and (6) are satisfied, the only combinations which are possible are

$$(\alpha) \text{ with } (a), \quad (\beta) \text{ with } (b) \quad \text{and} \quad (\gamma) \text{ with } (c),$$

which are exactly the combinations permitted by Euclid's Fifth Definition.

Hence, if Archimedes' Axiom be assumed, the satisfaction of the conditions (5) and (6) is a sufficient test for the proportion

$$A : B :: C : D.$$

If we take the Fifth Definition as translated by Heath and strike out the words

"alike exceed, are alike equal to, or",

the remaining words would embody conditions (5) and (6), though the meaning would not (I think) be very clear.

A similar demonstration to the above will shew that conditions (1) and (2) are a sufficient test of the proportion

$$A : B :: C : D.$$

These, with a like reservation as to the obscurity of the meaning, would be represented by striking out from Heath's translation of the definition the words

"are alike equal to, or alike fall short of."

7. It will now be shewn that (1) is equivalent to (6).

If (1) holds, then all values of $r$, $s$ which make $rA > sB$ also make $rC > sD$.

Now suppose that $r_1$, $s_1$ are such that $r_1 C < s_1 D$, it will be proved that $r_1 A < s_1 B$.

For, if not, either $r_1 A = s_1 B$ or $r_1 A > s_1 B$,

i.e. we have either (a) $r_1 C < s_1 D$ and $r_1 A = s_1 B$,

or (b) $r_1 C < s_1 D$ and $r_1 A > s_1 B$.

In case (a) $s_1 D - r_1 C$ is a magnitude of the same kind as $C$, and therefore by Archimedes' Axiom an integer $n$ exists such that

$$n(s_1 D - r_1 C) > C,$$

$$\therefore (nr_1 + 1) C < ns_1 D.$$

But since

$$r_1 A = s_1 B,$$

$$\therefore nr_1 A = ns_1 B,$$

$$\therefore (nr_1 + 1) A > ns_1 B.$$

Hence the integers $(nr_1 + 1)$ and $(ns_1)$ are such that

$$(nr_1 + 1) A > (ns_1) B, \text{ but } (nr_1 + 1) C < (ns_1) D,$$

which contradicts the condition (1).

In case (b) $r_1 A > s_1 B$, but $r_1 C < s_1 D$.

This also contradicts (1), for if $r_1 A > s_1 B$, then (1) requires that $r_1 C > s_1 D$.

Hence neither (a) nor (b) can hold, and therefore if $r_1 C < s_1 D$, then must $r_1 A < s_1 B$, and therefore condition (6) holds.

It remains to shew that if (6) holds, then (1) also holds.

If (6) holds, then all values of $r$, $s$ which make $rC < sD$ also make $rA < sB$.

Now suppose that $r_1$, $s_1$ are such that $r_1A > s_1B$, to prove that $r_1C > s_1D$.

For, if not, either $r_1C = s_1D$ or $r_1C < s_1D$,

i.e. we have either (c) $r_1A > s_1B$, $r_1C = s_1D$,

or (d) $r_1A > s_1B$, $r_1C < s_1D$.

In case (c) $r_1A - s_1B$ is a magnitude of the same kind as $B$, and therefore by Archimedes' Axiom an integer $n$ exists such that

$$n(r_1A - s_1B) > B,$$

$$\therefore\ nr_1A > (ns_1 + 1)B.$$

But since

$$r_1C = s_1D,$$

$$\therefore\ nr_1C = ns_1D,$$

$$\therefore\ nr_1C < (ns_1 + 1)D.$$

Hence the integers $(nr_1)$ and $(ns_1 + 1)$ are such that

$$(nr_1)C < (ns_1 + 1)D, \quad \text{but} \quad (nr_1)A > (ns_1 + 1)B,$$

which contradicts (6).

In case (d) $r_1C < s_1D$, $r_1A > s_1B$,

which also contradicts (6).

Hence neither (c) nor (d) can hold.

Hence if $r_1A > s_1B$, then must $r_1C > s_1D$, and therefore condition (1) holds.

Hence by the aid of Archimedes' Axiom it has been shewn that if (1) holds, then (6) holds; and if (6) holds, then (1) holds.

Hence (1) and (6) are equivalent.

By interchanging in the above proof $A$ with $C$ and $B$ with $D$, it follows that (2) and (5) are equivalent.

8. To sum up, the six conditions (1)—(6) involved in Euclid's Definition can be reduced by purely *logical* processes only to a smaller number in three ways, viz. to

(1), (3) and (5);

or to (2), (4) and (6);

or to (1), (2), (5) and (6).

If in addition to purely *logical* processes the truth of Archimedes' Axiom is assumed, then the six conditions can be reduced to two in the following ways, viz. to

(1) and (5);

or to (2) and (6);

or to (1) and (2);

or to (5) and (6).

Further, (1) is equivalent to (6), and (2) to (5).

The reduction to the pair (5) and (6) possesses certain advantages in dealing with some propositions over the other forms (see Arts. 11 and 14 below).

## V. The Propositions in the Fifth Book of Euclid. Their dependence on the Axiom of Archimedes.

9. The next step is to classify the propositions of the Fifth Book.

(i) The first group consists of propositions dealing with magnitudes and their multiples.

These are Nos. 1, 2, 3, 5 and 6.

With these should be included the following proposition:

> "If $A$, $B$, $C$ are magnitudes of the same kind, and if $A$ be greater than $B$, then integers $n$ and $t$ exist such that
>
> $$nA > tC > nB."$$

The proof of this forms the greater part of Prop. 8, and it depends on Euclid's Fourth Definition, so that Archimedes' Axiom is involved.

This proposition belongs properly to this first group, because it does not deal with ratios. In order to distinguish it from Prop. 8 I will refer to it in what follows as the principal part of Prop. 8.

The only place in the Fifth Book in which the Fourth Definition is used explicitly is in the proof of this principal part of Prop. 8.

(ii) The second group consists of propositions dealing with *Unequal* Ratios.

These are Nos. 8, 10 and 13.

The proofs of these depend on the Seventh Definition, the test for distinguishing between *Unequal* Ratios; whilst the proof of Prop. 8 (as has been already noted) requires also the Axiom of Archimedes.

The propositions in this group are used sometimes singly and sometimes all together in the proofs of Props. 9, 14, 16 and 18—25.

(iii) The third group consists of propositions dealing with *Equal* Ratios, which depend on the Fifth Definition and do not *necessarily* require the Axiom of Archimedes.

These are Nos. 4, 7, 11, 12, 15, 17 and 18.

Euclid's proofs of all of these except the last do not require the Axiom of Archimedes.

His proof of Prop. 18 assumes not only Prop. 8, and therefore the Axiom of Archimedes, but also the existence of a fourth proportional to three magnitudes, of which the first and second are of the same kind. Simson gave a proof free from either assumption. It is essentially the same as that in Art. 154 of the 2nd Edition of my *Euclid V. and VI.*

Another proof of Prop. 18 is given in Art. 14 below to illustrate the power of the proposition in Art. 6 above, but this assumes the Axiom of Archimedes, because that Axiom was employed in proving Art. 6.

(iv) The fourth group consists of propositions dealing with *Equal* Ratios which require both the Fifth Definition and the Axiom of Archimedes.

These are Nos. 9, 14, 16, 20—23.

Euclid's proofs of these propositions are made unnecessarily indirect and therefore difficult by his use of Props. 8, 10 and 13 in their proofs; thus bringing in the idea of *Unequal* Ratios to prove Properties of *Equal* Ratios.

It is shewn in the works mentioned in Art. 1 that these propositions can be proved by the aid of the Fifth Definition and the principal part only of Prop. 8; and it is shewn that then Props. 14, 20 and 21 can be treated as particular cases of Props. 16, 22 and 23, whilst the Euclidean method requires that Prop. 14 should be proved first as a stepping-stone for Prop. 16, and in like manner Prop. 20 for Prop. 22 and Prop. 21 for Prop. 23.

(v) The fifth group consists of propositions dealing with *Equal* Ratios which depend on propositions in the third and fourth groups.

These are Nos. 19, 24 and 25.

Euclid in his proofs employs only the propositions in the third and fourth groups. He does not make any direct use of the properties of *unequal* ratios with which the second group is concerned.

Inasmuch as proofs of the propositions in the third and fourth groups can be given which do not *necessarily* depend on the properties of *Unequal* Ratios, it is possible to regard the propositions in this fifth group as not depending *necessarily* on the properties of *Unequal* Ratios. They do however depend on the Axiom of Archimedes.

Euclid's proofs are I believe the simplest which can be given.

The proofs given in the works mentioned in Art. 1 of the propositions in the third and fourth groups are such that each proposition is deduced directly from the Fifth Definition, those in the fourth group requiring also the Axiom of Archimedes; but the proof of each proposition is independent of all the others.

In my second paper I attempted to obtain similar proofs of Props. 19, 24 and 25, but these, as will be seen on reference to Arts. 70—73 of that paper, are very complicated and indirect. I asked my friend Mr Rose-Innes if he could find something simpler. He has sent me those which now follow, Arts. 10—12. It is possible that no further simplification can be attained, but they are not as direct and the steps do not follow so automatically as those which I have given of Props. 16, 22 and 23 in my *Theory of Proportion*. I give also (Art. 14) a proof of Prop. 18 which will illustrate the power of the proposition in Art. 6.

EUC. V. 19.

10. Let $A$, $B$, $C$, $D$ be four magnitudes of the same kind such that

$$A : B = C : D, \text{ and } A > C, \quad B > D,$$

to prove $$A - C : B - D = A : B.$$

Take any multiple of $A$, say $rA$;

and any multiple of $B$, say $sB$.

There are three possibilities,

$$\text{(i) } rA < sB, \quad \text{(ii) } rA = sB, \quad \text{(iii) } rA > sB.$$

Consider (i). Since $rA < sB$,

therefore $sB - rA$ is a magnitude of the same kind as $A$, $B$, $C$, $D$, and since $A > C$, $A - C$ is a magnitude of the same kind as $A$, $B$, $C$, $D$.

Hence by Archimedes' Axiom an integer $p$ exists such that

$$p(A - C) > sB - rA,$$

$$\therefore\ pA > (sB - rA) + pC.$$

Hence* an integer $q$ exists such that

$$pA > q(sB - rA) > pC.$$

Now since $$pA > q(sB - rA),$$

$$\therefore\ (p + qr)A > (qs)B.$$

But $$A : B = C : D,$$

$$\therefore\ (p + qr)C > (qs)D,$$

$$\therefore\ pC + (qr)C > (qs)D \text{......(I)},$$

but $$q(sB - rA) > pC \text{......(II)},$$

$$\therefore\ q(sB - rA) + (qr)C > (qs)D \text{......from (I) and (II)},$$

$$\therefore\ sB - rA + rC > sD,$$

$$\therefore\ r(A - C) < s(B - D).$$

Hence if $$rA < sB,$$

then $$r(A - C) < s(B - D).$$

Consider (ii). Since $$rA = sB,$$

and $$A : B = C : D,$$

$$\therefore\ rC = sD,$$

$$\therefore\ r(A - C) = s(B - D).$$

Hence if $$rA = sB,$$

then $$r(A - C) = s(B - D).$$

Consider (iii). $$rA > sB.$$

Then since $$A > C,$$

therefore integers $p$, $q$ exist such that

$$pA > q(rA - sB) > pC.$$

Since $$pA > q(rA - sB),$$

$$(qs)B > (qr - p)A,$$

therefore we have *provided that* $$p < (qr),$$

$$(qs)D > (qr - p)C \quad \text{because} \quad A : B = C : D.$$

$$\therefore\ pC + (qs)D > (qr)C \text{......(III)}.$$

But $$q(rA - sB) > pC \text{......(IV)},$$

$$\therefore\ q(rA - sB) + (qs)D > (qr)C \text{......from (III) and (IV)},$$

$$\therefore\ rA - sB + sD > rC,$$

$$\therefore\ r(A - C) > s(B - D).$$

* If $X > Y + Z$, an integer $q$ exists such that $X > qY > Z$.

If however $p \not< (qr)$,

then it is still true that $(qs) D > (qr - p) C$,

for the left-hand side is positive and the right negative or zero,

$$\therefore pC + (qs) D > (qr) C.$$

But $q (rA - sB) > pC$,

$$\therefore q (rA - sB) + (qs) D > (qr) C,$$

$$\therefore rA - sB + sD > rC,$$

$$\therefore r (A - C) > s (B - D) \text{ as before.}$$

Hence if $rA > sB$,

then $r (A - C) > s (B - D)$.

It results from the three cases considered that

$$A - C : B - D = A : B.$$

EUC. V. 24.

[This demonstration illustrates the power of the theorem in Art. 6.]

11. To shew that if $A : C = X : Z$,

and if $B : C = Y : Z$,

then $A + B : C = X + Y : Z$.

(i) Let us suppose $r (A + B) < sC$.

Then $rA < r (A + B) < sC$.

Since $rA < sC$,

$sC - rA$ is a magnitude of the same kind as $A$, $B$, $C$.

Since $r (A + B) < sC$,

$$\therefore rB < sC - rA.$$

Consequently, assuming Archimedes' Axiom, integers $n$, $t$ exist such that

$$n (rB) < tC < n (sC - rA).$$

Since $(nr) B < tC$ and $B : C = Y : Z$,

$$\therefore (nr) Y < tZ.$$

Since $tC < n (sC - rA)$,

$$\therefore nrA + tC < nsC.$$

This involves $tC < nsC$,

$$\therefore t < ns,$$

$$\therefore nrA < (ns - t) C,$$

and since $A : C = X : Z$, $\therefore nrX < (ns - t) Z$,

but $nrY < tZ$,

$$\therefore nr (X + Y) < nsZ,$$

$$\therefore r (X + Y) < sZ.$$

Hence if $r (A + B) < sC$, then $r (X + Y) < sZ$.

(ii) If we suppose $r(X+Y) < sZ$ and re-write the above proof, interchanging $A$ with $X$, $B$ with $Y$ and $C$ with $Z$, we shall find that $r(A+B) < sC$.

From (i) and (ii) together we have

if $\qquad r(A+B) < sC, \text{ then } r(X+Y) < sZ,$

and if $\qquad r(X+Y) < sZ, \text{ then } r(A+B) < sC.$

Hence by the theorem of Art. 6

$$A+B : C = X+Y : Z.$$

### Euc. v. 25.

12. If $A$, $B$, $C$, $D$ be four magnitudes of the same kind in proportion, to shew that the sum of the greatest and least is greater than the sum of the other two.

Let $D$ be the least of the quantities.

Then $C-D$ is a magnitude of the same kind as $B$ and $D$, and $B > D$ because $D$ is the least of the quantities.

Hence, assuming Archimedes' Axiom, integers $p$, $q$ exist such that

$$pB > q(C-D) > pD,$$

$$\therefore\ q(C-D) > pD,$$

$$\therefore\ qC > (p+q)D.$$

But $\qquad A : B = C : D,$

$$\therefore\ qA > (p+q)B \ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \text{(I)},$$

but $\qquad pB > q(C-D) \ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \text{(II)},$

therefore from (I) and (II) $\qquad qA > q(C-D) + qB,$

$$\therefore\ A > C - D + B.$$

Since $D$ is the least, this inequality shews that

$A$ exceeds $B$ by $C-D$,

and $\qquad A$ exceeds $C$ by $B-D$.

Therefore $A$ is the greatest.

Moreover the inequality shews

$$A+D > B+C.$$

13. There is a certain resemblance between Mr Rose-Innes' proofs of Euc. v. 19 and Euc. v. 25. If we compare the inequalities marked (I) and (II) in v. 25 with those similarly marked in v. 19 or with those marked (III) and (IV) in v. 19 this resemblance will in part appear. It rests on the basis that either proposition, having first been proved independently, can be used to prove the other.

Euclid's proof of v. 25 depends on v. 19 and other propositions.

I will now shew how to use v. 25 and other propositions to prove v. 19.

Starting from $A : B = C : D$,

$\therefore\ rA : sB = rC : sD$ (v. 4).

Suppose, as in v. 19, that $A > C$, $B > D$.

These give $rA > rC$, $sB > sD$.

(i) If now $rA > sB$,

then since $rA : sB = rC : sD$,

$\therefore\ rC > sD$.

Hence $rA$ is the greatest and $sD$ the least of the four magnitudes in the proportion

$$rA : sB = rC : sD,$$

$$\therefore\ rA + sD > sB + rC \text{ by v. 25},$$

$$\therefore\ r(A - C) > s(B - D).$$

If therefore $rA > sB$, then $r(A - C) > s(B - D)$.

(ii) If $rA = sB$,

then $rC = sD$,

$\therefore\ r(A - C) = s(B - D)$.

If therefore $rA = sB$, then $r(A - C) = s(B - D)$.

(iii) If $rA < sB$,

then since $rA : sB = rC : sD$,

$\therefore\ rC < sD$:

we have also $rA > rC$,

and $sB > sD$,

therefore $sB$ is the greatest and $rC$ the least of the four magnitudes in the proportion

$$rA : sB = rC : sD,$$

$$\therefore\ sB + rC > rA + sD \text{ by Euc. v. 25},$$

$$\therefore\ r(A - C) < s(B - D).$$

If therefore $rA < sB$, then $r(A - C) < s(B - D)$.

From (i), (ii), (iii) it follows that

$$A - C : B - D = A : B.$$

## EUC. V. 18.

14. The following proof due to Mr Rose-Innes also illustrates the power of the proposition in Art. 6.

If $A : B = C : D$,

to prove that $A + B : B = C + D : D$.

Consider (i) the case $r(A + B) < sB$.

In this case $r$ must be less than $s$,

$$\therefore\ rA < (s-r)B,$$

$$\therefore\ rC < (s-r)D, \quad \because\ A : B = C : D,$$

$$\therefore\ r(C+D) < sD.$$

Hence if $\quad r(A+B) < sB,\ $ then $\ r(C+D) < sD.$

Consider (ii) the case $\quad r(C+D) < sD,$

then re-writing the above proof after interchanging $A$ with $C$ and $B$ with $D$, it follows that

$$r(A+B) < sB.$$

Hence from (i) and (ii) together,

if $\quad r(A+B) < sB,\ $ then $\ r(C+D) < sD,$

but if $\quad r(C+D) < sD,\ $ then $\ r(A+B) < sB.$

Hence by the theorem of Art. 6

$$A+B : B = C+D : D.$$

In the proof of this proposition Archimedes' Axiom has been assumed, because it was used in the demonstration of Art. 6.

This demonstration illustrates another point. Suppose that instead of using the conditions (5) and (6) of Art. 4 we had employed instead (1) and (2), so that we should have started with $r(A+B) > sB$.

In discussing this inequality we should have been obliged to discuss the three cases

$$r > s, \quad r = s, \quad r < s.$$

This justifies the remark at the end of Art. 8 that the use of conditions (5) and (6) may sometimes be more convenient than that of (1) and (2).

# VII. *The Invariants of the Halphenian Homographic Substitution—to which is appended some investigations concerning the Transformation of Differential Operators which present themselves in Invariant Theories.*

By Major P. A. MacMahon, F.R.S., Hon. Mem. Camb. Phil. Soc.

[*Received* 16 June, 1914.]

## INTRODUCTION.

This paper follows naturally the one published in these *Transactions* in 1908 under the title "The Operator Reciprocants of Sylvester's Theory of Reciprocants."

The particular object in view is the study of the invariant operators of the theories of invariants and reciprocants and the transformation of those operators. There is great advantage in adding operators to the invariant material dealt with. It was not at first recognized that the operators were effective because they themselves possessed invariant properties. The relations which establish those properties shew the exact conditions under which the operators are effective either as generators or annihilators. In certain cases homogeneity or isobarism or both may be necessary in the algebraic forms; in others the forms must possess properties in regard to other differential operators. The two simple substitutions of Sylvester and Halphen, both of period 2, suffice to disclose and elucidate the invariant properties and to discover the relations that exist between the two theories. What I have called the $b$ transformation, that was brought to light in the first paper, is herein further examined in regard to the special logarithmic case and two new transformations, the $h$ and the $s$, are discovered and examined. Transformations are shewn to exist which bring the seminvariants and pure reciprocant defining operators to the simplest possible forms, and shew instantaneously a complete system of ground forms in each case. The paper is divided into two sections—the first deals entirely with the Halphenian substitution, and the invariants, algebraic and operational, are exhibited in their categories. Attention may be directed to the symbolic method of Art. 8.

Section II treats of the transformation of linear operators in general, with special reference to the subject-matter of Section I.

## SECTION I.

### ON THE INVARIANTS OF THE HALPHENIAN HOMOGRAPHIC SUBSTITUTION

$$x = \frac{1}{X}, \quad y = \frac{Y}{X}.$$

1. If we consider the binary quantic

$$(a_0, a_1, a_2, \ldots)(u, v)^n,$$

and $y$ any function of $x$, we may suppose that

$$a_s = \frac{1}{(s+2)!} \frac{d^{s+2}y}{dx^{s+2}}.$$

If we now make the substitution

$$x = \frac{1}{X}, \quad y = \frac{Y}{X},$$

the invariants which are such that they are homogeneous and isobaric functions of

$$a_0, a_1, a_2, \ldots$$

are, as is well known, seminvariants of the binary quantic, and conversely every seminvariant is an invariant of the Halphenian substitution. They satisfy the well-known partial differential equation

$$\Omega_a = a_0 \partial_{a_1} + 2a_1 \partial_{a_2} + 3a_2 \partial_{a_3} + \ldots = 0.$$

In the paper communicated by me to the Cambridge Philosophical Society in 1908 I considered Sylvester's substitution

$$x = Y, \quad y = X,$$

the invariants of which, when homogeneous and isobaric functions of $a_0, a_1, a_2, \ldots$, have been called pure reciprocants. Such satisfy the differential equation

$$V_a = 4 \cdot \tfrac{1}{2} a_0^2 \partial_{a_1} + 5(a_0 a_1) \partial_{a_2} + 6(a_0 a_2 + \tfrac{1}{2} a_1^2) \partial_{a_3} + \ldots = 0.$$

Certain forms arise from both sets of substitutions and are thus both seminvariant and pure reciprocant functions of $a_0, a_1, a_2, \ldots$. Such are invariants for the general homographic substitution

$$x = \frac{\lambda_1 X + \mu_1 Y + \nu_1}{\lambda X + \mu Y + \nu}, \quad y = \frac{\lambda_2 X + \mu_2 Y + \nu_2}{\lambda X + \mu Y + \nu},$$

and have been named by Sylvester projective reciprocants and also principiants.

In fact principiants may be defined by the simultaneous partial differential equations

$$\Omega_a = 0, \quad V_a = 0.$$

Since moreover

$$\Omega_c = -V_b,$$

where

$$c_0 = \frac{1}{b_0},$$

$$2c_1 = -\frac{b_1}{b_0^3},$$

$$3c_2 = -\frac{b_2}{b_0^4} + \frac{2b_1^2}{b_0^5},$$

$$4c_3 = -\frac{b_3}{b_0} + \frac{5b_1b_2}{b_0} - \frac{5b_1^3}{b_0^2}.$$

..............................

$$b_0 = \frac{1}{c_0},$$

$$b_1 = -2\frac{c_1}{c_0^3},$$

$$b_2 = -\frac{3c_2}{c_0^4} + \frac{8c_1^2}{c_0^5},$$

$$b_3 = -\frac{4c_3}{c_0^5} + \frac{30c_1c_2}{c_0^6} - \frac{40c_1^3}{c_0^7}.$$

..............................

it follows that principiants are seminvariants which remain seminvariants after the substitution of

$$-2a_1, \quad -3a_2 + 8a_1^2, \quad -4a_3 + 30a_1a_2 - 40a_1^3, \ldots$$

for $a_1$, $a_2$, $a_3$, ... respectively,

and also after the substitution of

$$-\tfrac{1}{2}a_1, \quad -\tfrac{1}{3}a_2 + \tfrac{2}{3}a_1^2, \quad -\tfrac{1}{4}a_3 + \tfrac{5}{4}a_1a_2 - \tfrac{5}{4}a_1^3, \ldots$$

for the same quantities.

Ex. gr. The principiant

$$a_0^2a_3 - 3a_0a_1a_2 + 2a_1^3$$

is, to a power of $a_0$ *près*, merely multiplied by $-3$ by the first substitution and by $-\frac{1}{4}$ by the second, as may be readily verified.

2. In this Section of the paper I discuss the general invariant theory of the Halphenian substitution with particular reference to the invariant operators.

I write in a usual notation

$$\frac{dy}{dx} = t, \quad \frac{1}{s!}\frac{d^sy}{dx^s} = a_{s-2},$$

$$\frac{dY}{dX} = T, \quad \frac{1}{s!}\frac{d^sY}{dX^s} = A_{s-2}, \quad \text{where } s \not< 2.$$

Then, as shewn by Halphen,

$$t = -X.T - \frac{Y}{X},$$

$$a_0 = X^3A_0,$$

$$a_1 = -X^5\left(A_1 + \frac{A_0}{X}\right),$$

$$a_2 = X^7\left(A_2 + 2\frac{A_1}{X} + \frac{A_0}{X^2}\right).$$

..............................

$$a_s = (-)^s X^{2s+3}\left\{A_s + \binom{s}{1}\frac{A_{s-1}}{X} + \binom{s}{2}\frac{A_{s-2}}{X^2} + \ldots + \frac{A_0}{X^s}\right\}.$$

The Halphenian substitution, like that of Sylvester, has a period 2; it follows

(i) that the substitution may be given the symmetrical form

$$y = \frac{Y}{X^2}, \quad Y = \frac{y}{x^2};$$

(ii) that, in any relation involving the letters

$$y, x, t, a_0, a_1, \ldots \qquad Y, X, T, A_0, A_1, \ldots$$

we may interchange the small and capital letters.

We may further consider the operator symbols

$$\partial_y, \partial_x, \partial_t, \partial_{a_0}, \partial_{a_1}, \ldots \qquad \partial_Y, \partial_X, \partial_T, \partial_{A_0}, \partial_{A_1}, \ldots$$

and then if

$$\begin{aligned} &f(y, x, t, a_0, a_1, \ldots \partial_y, \partial_x, \partial_t, \partial_{a_0}, \partial_{a_1}, \ldots) \\ &= \phi(Y, X, T, A_0, A_1, \ldots \partial_Y, \partial_X, \partial_T, \partial_{A_0}, \partial_{A_1}, \ldots), \end{aligned}$$

the interchange of small and capital letters shews that

$$f(y, x, t, a_0, a_1, \ldots \partial_y, \partial_x, \partial_t, \partial_{a_0}, \partial_{a_1}, \ldots) \pm \phi(y, x, t, a_0, a_1, \ldots \partial_y, \partial_x, \partial_t, \partial_{a_0}, \partial_{a_1}, \ldots)$$

is an absolute operational invariant of the substitution of even or uneven order according as the upper or lower sign is taken.

In fact every symmetric function of

$$f(y, x, \ldots), \quad \phi(y, x, \ldots)$$

is an invariant of even order and every two-valued function of the two functions is an invariant of uneven order.

Thus the relations $\qquad a_0 = X^3 A_0, \quad x^{\frac{3}{2}} a_0 = X^{\frac{3}{2}} A_0,$

yield the invariants $\qquad (1 + x^3) a_0, \quad x^{\frac{3}{2}} a_0$ of even order

and $\qquad (1 - x^3) a_0$ of uneven order.

3. We assign to the letters $y, x, t, a_0, a_1, \ldots$ a certain degree $i$ and weight $w$ and deduce the characteristic $3i + 2w$ of the letter. We write $v$ for the characteristic and then we have the following scheme:

| | $y$ | $x$ | $t$ | $a_0$ | $a_1 \ldots$ | $a_s$ | $\partial_y$ | $\partial_x$ | $\partial_t$ | $\partial_{a_0}$ | $\partial_{a_1} \ldots$ | $\partial_{a_s}$ |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $i$ | 1 | 0 | 1 | 1 | 1 ... | 1 | −1 | 0 | −1 | −1 | −1 ... | −1 |
| $w$ | −2 | −1 | −1 | 0 | 1 ... | $s$ | 2 | 1 | 1 | 0 | −1 ... | $-s$ |
| $v$ | −1 | −2 | 1 | 3 | 5 ... | $2s+3$ | 1 | 2 | −1 | −3 | −5 ... | $-(2s+3)$ |

| | $Y$ | $X$ | $T$ | $A_0$ | $A_1 \ldots$ | $A_s$ | $\partial_Y$ | $\partial_X$ | $\partial_T$ | $\partial_{A_0}$ | $\partial_{A_1} \ldots$ | $\partial_{A_s}$ |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $i$ | 1 | 0 | 1 | 1 | 1 .. | 1 | −1 | 0 | −1 | −1 | −1 ... | −1 |
| $w$ | −1 | 1 | −2 | −3 | −4 ... | $-(s+3)$ | 1 | −1 | 2 | 3 | 4 ... | $s+3$ |
| $v$ | 1 | 2 | −1 | −3 | −5 ... | $-(2s+3)$ | −1 | −2 | 1 | 3 | 5 ... | $2s+3$ |

It will be remembered that for Sylvester's substitution the characteristic was $3i + w$; here it is $3i + 2w$.

The degree, weight and characteristic of a product of symbols are each of them formed by adding together the numbers which appertain to the symbols.

The investigation is much simplified by an accented notation; we multiply each letter or symbol (except $x$) by $x$ raised to the power of half of the characteristic of the letter or symbol; similarly for the capital letters. Thus we put

$$x^{-\frac{1}{2}}y = y', \qquad X^{-\frac{1}{2}}Y = Y',$$
$$x^{\frac{1}{2}}t = -t', \qquad X^{\frac{1}{2}}T = -T',$$
$$x^{\frac{3}{2}}a_0 = a_0', \qquad X^{\frac{3}{2}}A_0 = A_0',$$
$$x^{\frac{5}{2}}a_1 = -a_1', \qquad X^{\frac{5}{2}}A_1 = -A_1',$$
$$x^{\frac{1}{2}(2s+3)}a_s = (-)^s a_s', \qquad X^{\frac{1}{2}(2s+3)}A_s = (-)^s A_s'.$$

The change of sign will be noted. Also that what we have written

$$X^{\frac{1}{2}(2s+3)}A_s$$

is first of all $x^{-\frac{1}{2}(2s+3)}A_s$ in obedience to the law of formation.

The relations of Halphen become

$$y' = Y',$$
$$t' = -T' - Y',$$
$$a_0' = A_0',$$
$$a_1' = A_0' - A_1',$$
$$a_2' = A_0' - 2A_1' + A_2',$$
$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots$$
$$a_s' = A_0' - \binom{s}{1}A_1' + \binom{s}{2}A_2' - \ldots (-)^s A_s',$$

wherein there is no occurrence of $x$ or $X$.

For the accented letters we have the scheme:

| | $y'$ | $x'$ | $t'$ | $a_0'$ | $a_1'$ ... | $a_s'$ | $\partial_{y'}$ | $\partial_{x'}$ | $\partial_{t'}$ | $\partial_{a_0'}$ | $\partial_{a_1'}$ ... | $\partial_{a_s'}$ |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $i$ | 1 | 0 | 1 | 1 | 1 ... | 1 | $-1$ | 0 | $-1$ | $-1$ | $-1$ ... | $-1$ |
| $w$ | $-\frac{3}{2}$ | 0 | $-\frac{3}{2}$ | $-\frac{3}{2}$ | $-\frac{3}{2}$ ... | $-\frac{3}{2}$ | $\frac{3}{2}$ | 0 | $\frac{3}{2}$ | $\frac{3}{2}$ | $\frac{3}{2}$ ... | $\frac{3}{2}$ |
| $\nu$ | 0 | 0 | 0 | 0 | 0 ... | 0 | 0 | 0 | 0 | 0 | 0 ... | 0 |

| | $Y'$ | $X'$ | $T'$ | $A_0'$ | $A_1'$ ... | $A_s'$ | $\partial_{Y'}$ | $\partial_{X'}$ | $\partial_{T'}$ | $\partial_{A_0'}$ | $\partial_{A_1'}$ ... | $\partial_{A_s'}$ |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $i$ | 1 | 0 | 1 | 1 | 1 ... | 1 | $-1$ | 0 | $-1$ | $-1$ | $-1$ ... | $-1$ |
| $w$ | $-\frac{3}{2}$ | 0 | $-\frac{3}{2}$ | $-\frac{3}{2}$ | $-\frac{3}{2}$ ... | $-\frac{3}{2}$ | $\frac{3}{2}$ | 0 | $\frac{3}{2}$ | $\frac{3}{2}$ | $\frac{3}{2}$ ... | $\frac{3}{2}$ |
| $\nu$ | 0 | 0 | 0 | 0 | 0 ... | 0 | 0 | 0 | 0 | 0 | 0 ... | 0 |

so that each of the modified letters has the degree 1, the weight $-\frac{3}{2}$ and the characteristic zero; while each of the differential inverses has the degree $-1$, the weight $+\frac{3}{2}$ and the characteristic zero. Moreover every combination of accented letters and symbols has the characteristic zero.

It is to be observed that in every relation in accented notation it is permissible to interchange small and capital letters, both in symbols of quantity and in differential inverses.

The relations $a_1^{\backprime} = A_0^{\backprime} - A_1^{\backprime}, \quad a_2^{\backprime} = A_0^{\backprime} - 2A_1^{\backprime} + A_2^{\backprime}, \quad \&c.$

indicate that every function of $a_1^{\backprime},\ a_2^{\backprime}, \ldots$

is a function of differences of the quantities

$$A_0^{\backprime},\ A_1^{\backprime},\ A_2^{\backprime}, \ldots,$$

and thus satisfies the partial differential equation

$$\partial_{A_0^{\backprime}} + \partial_{A_1^{\backprime}} + \partial_{A_2^{\backprime}} + \ldots = 0.$$

In fact we will presently establish the formula of transformation

$$\partial_{a_0^{\backprime}} = \partial_{A_0^{\backprime}} + \partial_{A_1^{\backprime}} + \partial_{A_2^{\backprime}} + \ldots,$$

which is the analytical statement of the observed fact.

### *The Linear Invariants.*

4. Making use of the principle of invariance above set forth, the relations yield the absolute invariants of even order

$$y^{\backprime},$$
$$a_0^{\backprime},$$
$$a_0^{\backprime} - 2a_1^{\backprime} + 2a_2^{\backprime},$$
$$a_0^{\backprime} - 3a_1^{\backprime} + 3a_2^{\backprime},$$
$$a_0^{\backprime} - 4a_1^{\backprime} + 6a_2^{\backprime} - 4a_3^{\backprime} + 2a_4^{\backprime},$$
$$a_0^{\backprime} - 5a_1^{\backprime} + 10a_2^{\backprime} - 10a_3^{\backprime} + 5a_4^{\backprime},$$
$$\ldots\ldots\ldots\ldots\ldots\ldots$$

from which is derived the reduced set

| Accented | Unaccented |
|---|---|
| $y^{\backprime},$ | $x^{-\frac{1}{2}}y,$ |
| $a_0^{\backprime},$ | $x^{\frac{3}{2}}a_0,$ |
| $a_1^{\backprime} - a_2^{\backprime},$ | $-x^{\frac{5}{2}}a_1 - x^{\frac{7}{2}}a_2,$ |
| $a_2^{\backprime} - 2a_3^{\backprime} + a_4^{\backprime},$ | $x^{\frac{7}{2}}a_2 + 2x^{\frac{9}{2}}a_3 + x^{\frac{11}{2}}a_4,$ |
| $a_3^{\backprime} - 3a_4^{\backprime} + 3a_5^{\backprime} - a_6^{\backprime},$ | $-x^{\frac{9}{2}}a_3 - 3x^{\frac{11}{2}}a_4 - 3x^{\frac{13}{2}}a_5 - x^{\frac{15}{2}}a_6,$ |
| ………… | ………… |

Similarly we derive absolute invariants of uneven order which may be exhibited in the reduced forms

| Accented | Unaccented |
|---|---|
| $y^{\backprime} + 2t^{\backprime},$ | $x^{-\frac{1}{2}}y - 2x^{\frac{1}{2}}t,$ |
| $-a_0^{\backprime} + 2a_1^{\backprime},$ | $-x^{\frac{3}{2}}a_0 - 2x^{\frac{5}{2}}a_1,$ |
| $a_1^{\backprime} - 3a_2^{\backprime} + 2a_3^{\backprime},$ | $-x^{\frac{5}{2}}a_1 - 3x^{\frac{7}{2}}a_2 - 2x^{\frac{9}{2}}a_3,$ |
| $-a_2^{\backprime} + 4a_3^{\backprime} - 5a_4^{\backprime} + 2a_5^{\backprime},$ | $-x^{\frac{7}{2}}a_2 - 4x^{\frac{9}{2}}a_3 - 5x^{\frac{11}{2}}a_4 - 2x^{\frac{13}{2}}a_5,$ |
| $a_3^{\backprime} - 5a_4^{\backprime} + 9a_5^{\backprime} - 7a_6^{\backprime} + 2a_7^{\backprime},$ | $-x^{\frac{9}{2}}a_3 - 5x^{\frac{11}{2}}a_4 - 9x^{\frac{13}{2}}a_5 - 7x^{\frac{15}{2}}a_6 - 2x^{\frac{17}{2}}a_7,$ |
| ………… | ………… |

The general form here in its unaccented aspect may be taken as

$$x^{\frac{1}{2}(2s+3)}\left\{a_s+(s+2)\binom{s}{0}xa_{s+1}+\tfrac{1}{2}(s+3)\binom{s}{1}x^2a_{s+2}+\tfrac{1}{3}(s+4)\binom{s}{2}x^3a_{s+3}+\ldots+2x^{s+1}a_{2s+1}\right\}.$$

5. These invariants, linear in the quantities, may be exhibited in symbolic form by writing

$$a_s' = p^s, \quad A_s' = P^s,$$

and then

$$p^s = (1-P)^s,$$

$$p + P = 1;$$

and we derive the relations

$$p^s(1-p)^s = P^s(1-P)^s,$$

$$(1-2p)^s = (-)^s(1-2P)^s,$$

shewing that

$$p^s(1-p)^s(1-2p)^t = (-)^t P^s(1-P)^s(1-2P)^t,$$

which establishes that

$$p^s(1-p)^s(1-2p)^t$$

is an absolute invariant of even or uneven order according as $t$ is even or uneven.

Moreover the identity $\qquad (1-2p)^2 + 4p(1-p) = 1$

shews that we may exhibit the invariants of even order by forms

$$p^s(1-p)^s,$$

and those of uneven order by forms

$$p^s(1-p)^s(1-2p).$$

The sets of linear invariants above set forth are given in these symbolic forms by the successive integer values of $s$.

We have thus obtained the whole of the invariants which are linear in $a_0$, $a_1$, $a_2$, ... .

6. The theory of the invariants of higher orders in $a_0$, $a_1$, $a_2$, ... is very simple because we may take alternative symbols $p$, $q$, $r$, ... on the one hand and $P$, $Q$, $R$, ... on the other, where

$$p + P = q + Q = r + R = \ldots = 1,$$

and then

$$(p-q)^{s_{12}}(p-r)^{s_{13}}(q-r)^{s_{23}}\ldots = (-)^{s_{12}+s_{13}+s_{23}+\ldots}(P-Q)^{s_{12}}(P-R)^{s_{13}}(Q-R)^{s_{23}}\ldots,$$

and we thus obtain the whole series of invariants which present themselves in the invariant theory of an ordinary binary quantic. In fact we obtain all forms which satisfy the equations

$$\Omega_{a'} = a_0'\partial_{a_1'} + 2a_1'\partial_{a_2'} + 3a_2'\partial_{a_3'} + \ldots = 0.$$

Ex. gr. $\qquad (p-q)^2 = p^2q^0 - 2p^1q^1 + p^0q^2 = 2(a_0'a_2' - a_1'^2) = 2x^5(a_0a_2 - a_1^2),$

and

$$x^5(a_0a_2 - a_1^2) = X^5(A_0A_2 - A_1^2),$$

exhibiting the absolute invariance of $x^5(a_0a_2 - a_1^2)$.

The fact that the operator $\Omega_{a'}$ is itself an invariant under the substitution will shortly be established.

*The Invariant Operations.*

7. Let the functional equation $\qquad y = f(x)$

become by Halphen's substitution $\qquad Y = \phi(X)$.

Let $x$, $y$ receive simultaneous increments

$$\xi x, \quad \eta y,$$

and let the increments received by $X$, $Y$ in consequence be

$$\Xi X, \quad \mathrm{H} Y;$$

so that

$$y + \eta y = f(x + \xi x),$$

$$Y + \mathrm{H} Y = \phi(X + \Xi X),$$

where by reason of the substitution involved

$$(1 + \xi)(1 + \Xi) = 1,$$

$$1 + \mathrm{H} = \frac{1 + \eta}{1 + \xi} = (1 + \eta)(1 + \Xi);$$

and we deduce the relation

$$\frac{\eta}{\xi} + \frac{\mathrm{H}}{\Xi} = 1.$$

By Taylor's expansion we find

$$\eta y = t\xi x + a_0\xi^2 x^2 + a_1\xi^3 x^3 + \ldots,$$

$$\mathrm{H} Y = T\Xi X + A_0\Xi^2 X^2 + A_1\Xi^3 X^3 + \ldots.$$

Now, since

$$\frac{\eta}{\xi} + \frac{\mathrm{H}}{\Xi} = 1,$$

$$\frac{\eta}{\xi} - \frac{1}{2} = -\left(\frac{\mathrm{H}}{\Xi} - \frac{1}{2}\right).$$

Moreover

$$\frac{y}{x^{\frac{1}{2}}} = \frac{Y}{X^{\frac{1}{2}}},$$

so that

$$x^{\frac{1}{2}}\left(\frac{\eta y}{\xi x} - \frac{1}{2}\frac{y}{x}\right) = -X^{\frac{1}{2}}\left(\frac{\mathrm{H} Y}{\Xi X} - \frac{1}{2}\frac{Y}{X}\right).$$

Obtaining the expressions of $\frac{\eta y}{\xi x}$, $\frac{\mathrm{H} Y}{\Xi X}$, from Taylor's expansion above, and substituting herein we find

$$x^{\frac{1}{2}}\left(t - \frac{1}{2}\frac{y}{x} + a_0\xi x + a_1\xi^2 x^2 + \ldots\right) = -X^{\frac{1}{2}}\left(T - \frac{1}{2}\frac{Y}{X} + A_0\Xi X + A_1\Xi^2 X^2 + \ldots\right);$$

but

$$x^{\frac{1}{2}}\left(t - \frac{1}{2}\frac{y}{x}\right) = -X^{\frac{1}{2}}\left(T - \frac{1}{2}\frac{Y}{X}\right),$$

so that we are led to the relation

$$x^{\frac{1}{2}}(a_0\xi x + a_1\xi^2 x^2 + a_2\xi^3 x^3 + \ldots) = -X^{\frac{1}{2}}(A_0\Xi X + A_1\Xi^2 X^2 + A_2\Xi^3 X^3 + \ldots);$$

or proceeding to the accented notation

$$a_0'\xi - a_1'\xi^2 + a_2'\xi^3 - \ldots = -(A_0'\Xi - A_1'\Xi^2 + A_2'\Xi^3 - \ldots),$$

where

$$(1 + \xi)(1 + \Xi) = 1.$$

From Halphen's relations we deduce

$$\partial_y = X\partial_Y + \partial_T,$$

$$\partial_t = -\frac{1}{X}\partial_T,$$

$$(-)^s \partial_{a_s} = X^{-2s-3}\left(\partial_{A_s} - X^{-1}\partial_{A_{s+1}} + X^{-2}\partial_{A_{s+2}} - \dots\right);$$

and herein writing

$$x^{\frac{1}{2}}\partial_y = \partial_{y'}, \qquad X^{\frac{1}{2}}\partial_Y = \partial_{Y'},$$

$$x^{-\frac{1}{2}}\partial_t = -\partial_{t'}, \qquad X^{-\frac{1}{2}}\partial_T = -\partial_{T'},$$

$$x^{-\frac{1}{2}(2s+3)}\partial_{a_s} = (-)^s\partial_{a'_s}, \qquad X^{-\frac{1}{2}(2s+3)}\partial_{A_s} = (-)^s\partial_{A'_s},$$

we find

$$\partial_{y'} = \partial_{Y'} - \partial_{T'},$$

$$\partial_{t'} = -\partial_{T'},$$

$$(-)^s\partial_{a'_s} = \partial_{A'_s} + \binom{s+1}{1}\partial_{A'_{s+1}} + \binom{s+2}{2}\partial_{A'_{s+2}} + \dots.$$

8. To obtain these relations in symbolic form write

$$t' = a'_{-1}, \qquad y' = a'_{-2},$$

$$T' = A'_{-1}, \qquad Y' = A'_{-2},$$

and then put symbolically

$$(-)^s\partial_{a'_{s-1}} = k^s, \qquad (-)^{s+1}\partial_{A'_{s-1}} = K^s,$$

when we find that the relation

$$(-)^s\partial_{a'_s} = \partial_{A'_s} + \binom{s+1}{1}\partial_{A'_{s+1}} + \dots$$

may be written

$$k^{s+1} = (-)^{s+1}\left(\frac{K}{1+K}\right)^{s+1}.$$

Moreover, if $s = -1$, $k^0 = K^0$ yields $\partial_{t'} = -\partial_{T'}$;

if $s = -2$, $\frac{1}{k} = -\frac{1}{K} - k^0$ yields $\partial_{y'} = \partial_{Y'} - \partial_{T'}$.

The important observation is now made that the symbols $k$, $K$ are in fact related in the same manner as $\xi$, $\Xi$ for, from the relation

$$(1+\xi)(1+\Xi) = 1,$$

we at once deduce

$$\xi^{s+1} = (-)^{s+1}\left(\frac{\Xi}{1+\Xi}\right)^{s+1}.$$

Hence we may regard $\xi$ and $\Xi$ as symbols such that

$$\xi^s = (-)^s\partial_{a'_{s-1}}, \qquad \Xi^s = (-)^{s+1}\partial_{A'_{s-1}}.$$

This remarkable circumstance points to the important fact that in the relations

$$a'_0\xi - a'_1\xi^2 + a'_2\xi^3 - \dots = -\left(A'_0\Xi - A'_1\Xi^2 + A'_2\Xi^3 - \dots\right),$$

$$(1+\xi)(1+\Xi) = 1,$$

as well as in any rational integral relation connecting the quantities

$$\xi,\ y',\ t',\ a_0',\ a_1',\ \ldots$$

with

$$\Xi,\ Y',\ T',\ A_0',\ A_1',\ldots$$

we are at liberty to write

$$\xi^s = (-)^s \partial_{a'_{s-1}}, \qquad \Xi^s = (-)^{s+1} \partial_{A'_{s-1}},$$

where $s$ may have the values $-1,\ 0,\ 1,\ 2,\ 3,\ \ldots$

and

$$a'_{-2} = y',\ a'_{-1} = t',\ A'_{-2} = Y',\ A'_{-1} = T'.$$

From the relation $(1+\xi)(1+\Xi) = 1$ are obtained the useful relations

$$\xi^2 \partial_\xi = -\Xi^2 \partial_\Xi,$$

$$(1+\xi)\partial_\xi = -(1+\Xi)\partial_\Xi,$$

which indicate that, *quâ* invariant functions of $\xi$,

$$\xi^2 \partial_\xi \text{ and } (1+\xi)\partial_\xi$$

are invariant operations of uneven order.

When performed upon invariant functions of $\begin{matrix}\text{even}\\\text{uneven}\end{matrix}$ order they produce invariant functions of $\begin{matrix}\text{uneven}\\\text{even}\end{matrix}$ order.

*Invariant Operators of Zero-order in the Coefficients.*

9. These are all obtainable from the relation

$$(1+\xi)(1+\Xi) = 1,$$

for

$$\xi = -\frac{\Xi}{1+\Xi},$$

so that

$$\xi^s \pm (-)^s \left(\frac{\xi}{1+\xi}\right)^s$$

is an invariant in symbolic form, of even or uneven order, according as the upper or lower sign is taken.

We obtain the two series

| Even order | Uneven order |
|---|---|
| $\xi^0$, | $2\xi^{-1} + \xi^0$, |
| $\xi - \frac{\xi}{1+\xi}$, | $\xi + \frac{\xi}{1+\xi}$, |
| $\xi^2 + \frac{\xi^2}{(1+\xi)^2}$, | $\xi^2 - \frac{\xi^2}{(1+\xi)^2}$, |
| $\xi^3 - \frac{\xi^3}{(1+\xi)^3}$, | $\xi^3 + \frac{\xi^3}{(1+\xi)^3}$, |
| $\vdots$ | $\vdots$ |

Observe that by reason of the difference in sign, for a given value of $s$, in the relations

$$\xi^s = (-)^s \partial_{a'_{s-1}}, \qquad \Xi^s = (-)^{s+1} \partial_{A'_{s-1}}$$

a symbolic form of $\begin{matrix}\text{even}\\\text{uneven}\end{matrix}$ order yields an unsymbolic form of $\begin{matrix}\text{uneven}\\\text{even}\end{matrix}$ order. We thus obtain the invariant operations

Uneven order

$$\partial_{t'},$$

$$\partial_{a_1'} + \partial_{a_2'} + \partial_{a_3'} + \dots,$$

$$2\partial_{a_1'} + 2\partial_{a_2'} + 3\partial_{a_3'} + 4\partial_{a_4'} + \dots,$$

$$3\partial_{a_3'} + 6\partial_{a_4'} + 10\partial_{a_5'} + 15\partial_{a_6'} + \dots,$$

$$2\partial_{a_3'} + 4\partial_{a_4'} + 10\partial_{a_5'} + 20\partial_{a_6'} + \dots,$$

$$2\partial_{a'_{2s-1}} + \binom{2s}{1}\partial_{a'_{2s}} + \binom{2s+1}{2}\partial_{a'_{2s+1}} + \binom{2s+2}{3}\partial_{a'_{2s+2}} + \dots.$$

$$\binom{2s+1}{1}\partial_{a'_{2s+1}} + \binom{2s+2}{2}\partial_{a'_{2s+2}} + \binom{2s+3}{3}\partial_{a'_{2s+3}} + \dots.$$

From these we may derive a reduced set

$$\partial_{t'},$$

$$\partial_{a_1'} + \partial_{a_2'} + \partial_{a_3'} + \dots,$$

$$\partial_{a_3'} + 2\partial_{a_4'} + 3\partial_{a_5'} + \dots,$$

.........................

and, for $s \geqslant 0$,

$$\binom{2s+3}{3}\partial_{a'_{2s+3}} + 2\binom{2s+4}{4}\partial_{a'_{2s+4}} + 3\binom{2s+5}{5}\partial_{a'_{2s+5}} + \dots.$$

The unaccented expression of the operations is

$$x^{-\frac{1}{2}}\partial_t,$$

$$x^{-\frac{5}{2}}(\partial_{a_1} - x^{-1}\partial_{a_2} + x^{-2}\partial_{a_3} - \dots),$$

$$x^{-\frac{9}{2}}(\partial_{a_3} - 2x^{-1}\partial_{a_4} + 3x^{-2}\partial_{a_5} - \dots),$$

....................................

In obtaining these it must be remembered that the characteristics of $a_s$ and $\partial_{a_s}$ are the same numerically but differ in sign.

Even order

$$2\partial_{y'} - \partial_{t'},$$

$$2\partial_{a_0'} + \partial_{a_1'} + \partial_{a_2'} + \dots,$$

$$2\partial_{a_2'} + 3\partial_{a_3'} + 4\partial_{a_4'} + \dots,$$

$$2\partial_{a_2'} + 3\partial_{a_3'} + 6\partial_{a_4'} + 10\partial_{a_5'} + \dots,$$

....................................

$$\binom{2s}{1}\partial_{a'_{2s}} + \binom{2s+1}{2}\partial_{a'_{2s+1}} + \binom{2s+2}{3}\partial_{a'_{2s+2}} + \dots,$$

$$2\partial_{a'_{2s}} + \binom{2s+1}{1}\partial_{a'_{2s+1}} + \binom{2s+2}{2}\partial_{a'_{2s+2}} + \dots,$$

for $s \geqslant 1$.

The reduced set is

$$2\partial_{y'} - \partial_{t'},$$
$$2\partial_{a_0'} + \partial_{a_1'} + \partial_{a_2'} + \ldots,$$
$$2\partial_{a_2'} + 3\partial_{a_3'} + 4\partial_{a_4'} + \ldots,$$
$$2\partial_{a_4'} + 5\partial_{a_5'} + 9\partial_{a_6'} + \ldots,$$
$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots$$

of which the unsymbolic forms are

$$2x^{\frac{1}{2}}\partial_y + x^{-\frac{1}{2}}\partial_t,$$
$$x^{-\frac{3}{2}}(2\partial_{a_0} - x^{-1}\partial_{a_1} + x^{-2}\partial_{a_2} - \ldots),$$
$$x^{-\frac{7}{2}}(2\partial_{a_2} - 3x^{-1}\partial_{a_3} + 4x^{-2}\partial_{a_4} - \ldots),$$
$$x^{-\frac{11}{2}}(2\partial_{a_4} - 5x^{-1}\partial_{a_5} + 9x^{-2}\partial_{a_6} - \ldots),$$
$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots$$

To verify the first of these we have

$$2x^{\frac{1}{2}}\partial_y + x^{-\frac{1}{2}}\partial_t = 2X^{-\frac{1}{2}}(X\partial_Y + \partial_T) + X^{\frac{1}{2}}\left(-\frac{1}{X}\partial_T\right) = 2X^{\frac{1}{2}}\partial_Y + X^{-\frac{1}{2}}\partial_T.$$

We may also note that $\qquad \dfrac{\xi^2}{1+\xi} = \dfrac{\Xi^2}{1+\Xi},$

so that $\dfrac{\xi^{2s}}{(1+\xi)^s}$ is an invariant operator in symbolic form which denotes in unsymbolic form an operator of uneven order; it is

$$\partial_{a'_{2s-1}} + \binom{s}{1}\partial_{a'_{2s}} + \binom{s+1}{2}\partial_{a'_{2s+1}} + \ldots,$$

and in unaccented form

$$x^{-2s-\frac{1}{2}}\left\{\partial_{a_{2s-1}} - \binom{s}{1}x^{-1}\partial_{a_{2s}} + \binom{s+1}{2}x^{-2}\partial_{a_{2s+1}} - \ldots\right\}.$$

*Operators of the First Degree in the Coefficients.*

10. We have before us the two relations

$$p + P = 1,$$
$$(1+\xi)(1+\Xi) = 1,$$

and the established relation

$$(a_0'\xi - a_1'\xi^2 + a_2'\xi^3 - \ldots) = -(A_0'\Xi - A_1'\Xi^2 + A_2'\Xi^3 - \ldots).$$

It yields the relation

$$a_0'\partial_{a_0'} + a_1'\partial_{a_1'} + a_2'\partial_{a_2'} + \ldots = A_0'\partial_{A_0'} + A_1'\partial_{A_1'} + A_2'\partial_{A_2'} + \ldots,$$

or, as we shall write it, $\qquad I_{a'} = I_{A'}.$

This gives us the invariant of even order

$$I_a = a_0'\partial_{a_0'} + a_1'\partial_{a_1'} + a_2'\partial_{a_2'} + \ldots.$$

In unaccented form this is

$$I_a = a_0\partial_{a_0} + a_1\partial_{a_1} + \ldots = A_0\partial_{A_0} + A_1\partial_{A_1} + \ldots = I_A.$$

To obtain the operator in a form which is wholly symbolic we write

$$a_s' = p^s, \qquad A_s' = P^s,$$

and we then have

$$\frac{\xi}{1+p\xi} = -\frac{\Xi}{1+P\Xi},$$

so that $\frac{\xi}{1+p\xi}$ is the sought symbolic form.

This can be obtained directly from the relations

$$p + P = 1, \quad (1+\xi)(1+\Xi) = 1,$$

for these may be written

$$p - \tfrac{1}{2} = -(P - \tfrac{1}{2}),$$

$$\frac{1}{\xi} + \tfrac{1}{2} = -\left(\frac{1}{\Xi} + \tfrac{1}{2}\right),$$

and by addition we get

$$p + \frac{1}{\xi} = -\left(P + \frac{1}{\Xi}\right),$$

or

$$\frac{\xi}{1+p\xi} = -\frac{\Xi}{1+P\Xi}.$$

Since

$$\left(\frac{\xi}{1+p\xi}\right)^s = (-)^s\left(\frac{\Xi}{1+P\Xi}\right)^s$$

we find by expansion and interpretation

$$a_0'\partial_{a'_{s-1}} + \binom{s}{1} a_1'\partial_{a'_s} + \binom{s+1}{2} a_2'\partial_{a'_{s+1}} + \dots$$

$$= (-)^{s+1}\left\{A_0'\partial_{A'_{s-1}} + \binom{s}{1} A_1'\partial_{A'_s} + \binom{s+1}{2} A_2'\partial_{A'_{s+1}} + \dots\right\}.$$

This in unaccented notation is

$$x^{-s+1}\left\{a_0\partial_{a_{s-1}} + \binom{s}{1} a_1\partial_{a_s} + \binom{s+1}{2} a_2\partial_{a_{s+1}} + \dots\right\}$$

$$= (-)^{s+1}X^{-s+1}\left\{A_0\partial_{A_{s-1}} + \binom{s}{1} A_1\partial_{A_s} + \binom{s+1}{2} A_2\partial_{A_{s+1}} + \dots\right\},$$

establishing that

$$x^{-s+1}\left\{a_0\partial_{a_{s-1}} + \binom{s}{1} a_1\partial_{a_s} + \binom{s+1}{2} a_2\partial_{a_{s+1}} + \dots\right\}$$

is an absolute invariant of even or uneven order according as $s$ is uneven or even.

For $s = 2$, we find

$$a_0'\partial_{a_1'} + 2a_1'\partial_{a_2'} + 3a_2'\partial_{a_3'} + \dots = -A_0'\partial_{A_1'} + 2A_1'\partial_{A_2'} + 3A_2'\partial_{A_3'} + \dots$$

or in Sylvester's notation

$$\Omega_{a'} = -\Omega_{A'},$$

and

$$x^{-1}\Omega_a = -X^{-1}\Omega_A.$$

$\Omega_a$ is the operator which causes all seminvariants of a binary quantic to vanish. We see that, in this theory, $x^{-1}\Omega_a$ is an absolute invariant of uneven order. The operation either causes the invariant operand to vanish or produces an invariant of contrary order.

The absolute invariance of $I_a$ defined above clearly shews that every invariant is homogeneous in the letters $a_0, a_1, a_2, \ldots$.

The above series of invariant operators, linear in $a_0', a_1', a_2', \ldots$, can also be obtained by repeated operation of $\xi^2\partial_\xi$ upon $\xi/1+p\xi$. The former operation gives invariants because

$$\xi^2\partial_\xi = -\Xi^2\partial_\Xi.$$

The relation

$$\frac{1+p\xi}{\xi} = -\frac{1+P\Xi}{\Xi}$$

yields

$$a_1'\partial_{t'} - a_0'\partial_{y'} = A_1'\partial_{T'} - A_0'\partial_{Y'},$$

equivalent to

$$x^2(a_1\partial_t - a_0\partial_y) = X^2(A_1\partial_T - A_0\partial_Y),$$

establishing that

$$x^2(a_1\partial_t - a_0\partial_y)$$

is an absolute invariant of even order.

The operation of $(1+\xi)\partial_\xi$ upon $\dfrac{\xi}{1+p\xi}$ gives

$$\frac{1+\xi}{(1+p\xi)^2},$$

and yields the invariant of uneven order

$$a_0'\partial_{t'} + (2a_1' - a_0')\partial_{a_0'} + (3a_2' - 2a_1')\partial_{a_1'} + (4a_3' - 3a_2')\partial_{a_2'} + \ldots,$$

equivalent to

$$xa_0\partial_t + (2xa_1 + a_0)\partial_{a_0} + (3xa_2 + 2a_1)\partial_{a_1} + (4xa_3 + 3a_2)\partial_{a_2} + \ldots.$$

The operation may be repeated indefinitely.

*The operator* $W_{a'} = a_1'\partial_{a_1'} + 2a_2'\partial_{a_2'} + 3a_3'\partial_{a_3'} + \ldots$.

11. The symbolic form is $\dfrac{p\xi^2}{(1+p\xi)^2}$ equal to $(1-P)\dfrac{\Xi^2}{(1+P\Xi)^2}$,

shewing us that

$$W_{a'} = -\Omega_{A'} + W_{A'}:$$

and establishing that

$$2W_{a'} - \Omega_{a'}$$

is an absolute invariant of even order.

It is equivalent to $2W_a - x^{-1}\Omega_a$ because $W_{a'} = W_a$.

We see from this result that $W_a$ is not an invariant, so that every invariant is not an isobaric function of $a_0, a_1, a_2, \ldots$; but that this is the case, exceptionally, when the operand satisfies the equation $\Omega_a = 0$.

12. Two operators now present themselves for examination, viz.

| Accented Notation | Unaccented Notation |
|---|---|
| $w_{t'} = -t'\partial_{t'} + W_{a'}$, | $w_t = -t\partial_t + W_a$, |
| $w_{y'} = -2y'\partial_{y'} + W_{t'}$, | $w_y = -2y\partial_y + W_t$. |

Since

$$t' = -T' - Y', \quad \partial_{t'} = -\partial_{T'},$$

we find

$$W_{t'} = -(T' + Y')\partial_{T'} - \Omega_{A'} + W_{A'} = W_{T'} - Y'\partial_{T'} - \Omega_{A'}:$$

so that $2W_{t'} - y'\partial_{t'} - \Omega_{a'}$ is an invariant of even order.

This is equivalent to $$2W_t + x^{-1}y\partial_t - x^{-1}\Omega_a.$$

Again, $$W_{y'} = -2Y'(\partial_{T'} - \partial_{T'}) + W_{T'} - Y'\partial_{T'} - \Omega_{A'} = W_{Y'} + Y'\partial_{T'} - \Omega_{A'},$$

so that $$2W_{y'} + y'\partial_{t'} - \Omega_a,$$

equivalent to $$2W_y - x^{-1}y\,\partial_t - x^{-1}\Omega_a,$$

is an invariant of even order.

*The operators*

$$I_{t'} = t'\partial_{t'} + I_{a'}, \qquad \textit{equivalent to} \qquad I_t = t\partial_t + I_a,$$

$$I_{y'} = y'\partial_{y'} + I_{t'}, \qquad\qquad\qquad I_y = y\partial_y + I_t.$$

13. We obtain easily

$$I_{t'} = I_{T'} + Y'\partial_{T'}, \qquad \text{equivalent to} \qquad I_t = I_T - X^{-1}Y\partial_T,$$

$$I_{y'} = I_{Y'}, \qquad\qquad\qquad I_y = I_Y,$$

establishing the invariants of even order

$$2I_t - \frac{y}{x}\partial_t,$$

$$I_y.$$

The fact that $I_y$ is an invariant shews that every invariant is a homogeneous function of $y, a_0, a_1, a_2, \ldots$; but since $I_t$ is not an invariant, every invariant is not a homogeneous function of $t, a_0, a_1, a_2, \ldots$.

*Transformation of a General Multilinear Operator.*

14. Write

$$\frac{1}{m}(a_0' + a_1'u + a_2'u^2 + \ldots)^m = a'_{m0} + a'_{m1}u + a'_{m2}u^2 + \ldots,$$

and consider the operator

$$\mu a'_{m0}\partial_{a'_m} + (\mu + \nu)a'_{m1}\partial_{a'_{n+1}} + (\mu + 2\nu)a'_{m2}\partial_{a'_{n+2}} + \ldots.$$

Herein $\mu$, $\nu$ may be any real numerical magnitudes, zero included; $m$ also may be any real numerical magnitude, but will usually be a positive or negative integer and more usually still a positive integer. The zero value of $m$ has been shewn by Hammond to be connected with the function

$$\log(a_0' + a_1'u + a_2'u^2 + \ldots).$$

and will be considered later.

$n$ may be taken to be zero or any positive integer.

The operator under consideration may be briefly written

$$(\mu, \nu;\ m, n)_{a'} = (-)^n x^{\frac{1}{2}(3m-2n-3)}(\mu, \nu;\ m, n)_a,$$

a relation indicating the accented and unaccented forms respectively.

Write $$a_\xi' = a_0'\xi - a_1'\xi^2 + a_2'\xi^3 - \ldots,$$

$$A_\Xi' = A_0'\Xi - A_1'\Xi^2 + A_2'\Xi^3 - \ldots,$$

so that, as has been shewn above,

$$a_\xi' = -A_\Xi'.$$

In the symbolism explained above,

$$(-)^{n+1}(\mu, \nu; m, n)_{a'}$$

$$= \left(\frac{\mu}{m} - \nu\right) \xi^{n-m+1} a_\xi'^m + \nu \xi^{n-m+2} a_\xi'^{m-1} \partial_\xi a_\xi',$$

and to carry out the transformation we have the relations

$$a_\xi' = - A_\Xi',$$

$$\xi = - \frac{\Xi}{1+\Xi},$$

$$\xi^2 \partial_\xi = - \Xi^2 \partial_\Xi.$$

We then find $\qquad (-)^{n+1}(\mu, \nu; m, n)_{a'}$

$$= (-)^{n+1}\left\{\left(\frac{\mu}{m} - \nu\right) \Xi^{n-m+1}(1+\Xi)^{-n+m-1} A_\Xi'^m + \nu \Xi^{n-m+2}(1+\Xi)^{-n+m} A_\Xi'^{m-1} \partial_\Xi A_\Xi'\right\}.$$

Write now $\qquad (1+\Xi)^k = \Sigma C_{s,k} \Xi^s,$

so that $C_{s,k} = \binom{k}{s}$ if $k$ be a positive integer $\geqslant s$.

We then have $\qquad (-)^{n+1}(\mu, \nu; m, n)_{a'}$

$$= \sum_s \left(\frac{\mu}{m} - \nu\right) C_{s,\, -n+m-1} \Xi^{n-m+s+1} A_\Xi'^m$$

$$+ \sum_s \nu C_{s,\, -n+m} \Xi^{n-m+s+2} A_\Xi'^{m-1} \partial_\Xi A_\Xi'.$$

Comparing the general term herein with the symbolic form of

$$(-)^{n_1+1}(\mu_1, \nu_1; m_1, n_1)_{A'},$$

viz. $\qquad \left(\frac{\mu_1}{m_1} - \nu_1\right) \Xi^{n_1-m_1+1} A_\Xi'^{m_1} + \nu_1 \Xi^{n_1-m_1+2} A_\Xi'^{m_1-1} \partial_\Xi A_\Xi',$

we find

$$\mu_1 = m\nu C_{s,\, -n+m} + (\mu - m\nu) C_{s,\, -n+m-1},$$

$$\nu_1 = \nu C_{s,\, -n+m},$$

$$m_1 = m,$$

$$n_1 = n + s.$$

Hence $\qquad (\mu, \nu; m, n)_{a'}$

$$= \sum_{s=0} (-)^{n+s} \{\mu C_{s,\, -n+m-1} + m\nu (C_{s,\, -n+m} - C_{s,\, -n+m-1}),\, \nu C_{s,\, -n+m};\, m,\, n+s\}_{A'},$$

shewing that the transformation produces a sum of multilinear operators of the same general form and of the same degree in the coefficients.

15. Leaving out of consideration for the moment a zero value of $m$, we find that in some cases the transformation produces a single operator. There are two cases.

CASE I. If $\qquad \mu = m\nu, \quad -n + m = 0,$

$$(n, 1; n, n)_{a'} = (-)^n (n, 1; n, n)_{A'},$$

or

$$x^{\frac{1}{2}(n-3)} (n, 1; n, n)_a = (-)^n X^{\frac{1}{2}(n-3)} (n, 1; n, n)_A,$$

establishing that $$x^{\frac{1}{2}(n-3)}(n, 1; n, n)_a$$

is an absolute invariant of even or uneven order according as $n$ is even or uneven.

For $n=1$ we have

$$\Omega_{a'} = -\Omega_{A'} \text{ equivalent to } x^{-1}\Omega_a = -X^{-1}\Omega_A,$$

a relation already met with.

For $n=2$, $$(2, 1; 2, 2)_{a'} = (2, 1; 2, 2)_{A'},$$

which we write $$J_{a'} = J_{A'},$$

equivalent to $$x^{-\frac{1}{2}}J_a = X^{-\frac{1}{2}}J_A,$$

where $J_a$ is the important operator which generates pure reciprocants from pure reciprocants and seminvariants from seminvariants, and causes the vanishing of forms which are both pure reciprocants and seminvariants. The proof of this was given by me in the paper* and communicated by me to Sylvester. The latter incorporated the theorem into his Lectures on Reciprocants delivered before the University of Oxford.

In the present theory of the Halphenian transformation we see that

$$x^{-\frac{1}{2}}J_a$$

is an absolute invariant of even order.

CASE II. Let $$\nu = 0, \quad -n+m-1=0.$$

Then $$(1, 0; n+1, n)_{a'} = (-)^n(1, 0; n+1, n)_{A'},$$

equivalent to $$x^{\frac{1}{2}n}(1, 0; n+1, n)_a = (-)^n X^{\frac{1}{2}n}(1, 0; n+1, n)_A,$$

a relation already met with in the form

$$a_{\xi}^{\backslash n+1} = (-)^{n+1} A_{\Xi}^{\backslash n+1}.$$

There are no other cases.

16. We can also specify the conditions under which the transformation produces a sum of any given number of operators. Whatever the given number there are invariably two cases. Thus

$$(-)^n(\mu, \nu; n+1, n)_{a'}$$
$$= (\mu, \nu; n+1, n)_{A'} - \nu(n+1, 1; n+1, n+1)_{A'},$$

equivalent to $$x^{\frac{1}{2}n}(\mu, \nu; n+1, n)_a$$
$$= (-)^n\{X^{\frac{1}{2}n}(\mu, \nu; n+1, n)_A + X^{\frac{1}{2}(n-2)}(n+1, 1; n+1, n+1)_A\},$$

yielding the special cases

$$\mu I_{a'} + \nu W_{a'} = \mu I_{A'} + \nu W_{A'} - \nu\Omega_{A'},$$
$$V_{a'} = -V_{A'} + J_{A'};$$

establishing the invariant $$2\mu I_{a'} + \nu(2W_{a'} - \Omega_{a'})$$

of even order, and the invariant $$2V_{a'} - J_{a'}$$

of uneven order.

* *Proc. Lond. Math. Soc.* Vol. XVIII. p. 75, 1886.

Also
$$(-)^n (1, 0;\ n+2, n)_{a'}$$
$$= (1, 0;\ n+2, n)_{A'} - (1, 0;\ n+2, n+1)_{A'}.$$

Again for three operators on the dexter
$$(-)^n (\mu, \nu;\ n+2, n)_{a'}$$
$$= (\mu, \nu;\ n+2, n)_{A'} - \{\mu + (n+2)\nu, 2\nu;\ n+2, n+1\}_{A'} + \nu (n+2, 1;\ n+2, n+2)_{A'};$$
$$(-)^n (1, 0;\ n+3, n)_{a'}$$
$$= (1, 0;\ n+3, n)_{A'} - 2(1, 0;\ n+3, n+1)_{A'} + (1, 0;\ n+3, n+2)_{A'}.$$

Observe that if we multiply up by $m$ in the general formula and then put $m=0$, $\mu=1$, $\nu=0$ we obtain a formula already reached, viz.
$$(-)^n \partial_{a'_n} = \partial_{A'_n} + \binom{n+1}{1} \partial_{A'_{n+1}} + \binom{n+2}{2} \partial_{A'_{n+2}} + \dots.$$

*The logarithmic case.*

17. The case corresponding to $m=0$ will be best understood from a consideration of the two operators
$$S_a = \frac{a_1}{a_0} \partial_{a_0} + \frac{2a_0a_2 - a_1^2}{a_0^2} \partial_{a_1} + \frac{3a_0^2a_3 - 3a_0a_1a_2 + a_1^3}{a_0^3} \partial_{a_2} + \dots,$$
$$R_a = \frac{a_1}{a_0} \partial_{a_1} + \frac{2a_0a_2 - a_1^2}{a_0^2} \partial_{a_2} + \frac{3a_0^2a_3 - 3a_0a_1a_2 + a_1^3}{a_0^3} \partial_{a_3} + \dots.$$

We observe that
$$\log\left(1 + \frac{a_1}{a_0} u + \frac{a_2}{a_0} u^2 + \frac{a_3}{a_0} u^3 + \dots\right)$$
$$= \frac{a_1}{a_0} u + \frac{1}{2} \frac{2a_0a_2 - a_1^2}{a_0^2} u^2 + \frac{1}{3} \frac{3a_0^2a_3 - 3a_0a_1a_2 + a_1^3}{a_0^3} u^3 + \dots;$$
so that putting $a_0 + a_1 u + a_2 u^2 + \dots = U$, we may briefly denote
$$S_a \text{ by } \left(0, 1;\ \log\frac{U}{a_0}, -1\right)_a,$$
and
$$R_a \text{ by } \left(0, 1;\ \log\frac{U}{a_0}, 0\right)_a,$$
just as we might have written
$$(\mu, \nu;\ m, n)_a \text{ in the notation } \left(\mu, \nu;\ \frac{1}{m} U^m, n\right)_a.$$

Proceeding to accented forms we find
$$S_a = -x^{\frac{1}{2}} S_{a'},$$
$$R_a = x^{\frac{3}{2}} R_{a'}.$$

In symbolic forms
$$S_{a'} = -\frac{a_1'}{a_0'} \xi + \frac{2a_0'a_2' - a_1'^2}{a_0'^2} \xi^2 - \frac{3a_0'^2a_3' - 3a_0'a_1'a_2' + a_1'^3}{a_0'^3} \xi^3 + \dots,$$
$$R_{a'} = -\xi S_{a'}.$$

Now, putting as on a previous page,
$$a_\xi' = a_0'\xi - a_1'\xi^2 + a_2'\xi^3 - \dots.$$

we have $$\log a_\xi' = \log(a_0'\xi) + \log\left(1 - \frac{a_1'}{a_0'}\xi + \frac{a_2'}{a_0'}\xi^2 - \frac{a_3'}{a_0'}\xi^3 + \dots\right),$$

and differentiating with regard to $\xi$,

$$\partial_\xi \log a_\xi' = \xi^{-1} + \left\{-\frac{a_1'}{a_0'} + \frac{2a_0'a_2' - a_1'^2}{a_0'^2}\xi - \frac{3a_0'^2a_3' - 3a_0'a_1'a_2' + a_1'^3}{a_0'^3}\xi^2 + \dots\right\},$$

or $$\xi\partial_\xi \log a_\xi' = \xi^0 + S_{a'},$$
$$\xi^2\partial_\xi \log a_\xi' = \xi - R_{a'}.$$

Hence $$S_{a'} = \xi\partial_\xi \log a_\xi' - \xi^0 = \frac{\xi}{a_\xi'}\partial_\xi a_\xi' - \xi^0,$$

$$R_{a'} = -\xi^2\partial_\xi \log a_\xi' + \xi = -\frac{\xi^2}{a_\xi'}\partial_\xi a_\xi' + \xi.$$

Transforming from $\xi$ to $\Xi$,

$$S_a = \frac{(\Xi + \Xi^2)\partial_\Xi A_\Xi'}{A_\Xi'} - 1,$$

$$R_{a'} = \frac{\Xi^2\partial_\Xi A_\Xi'}{A_\Xi'} - \frac{\Xi}{1+\Xi};$$

from which $$S_{a'} - R_{a'} = \frac{\Xi\partial_\Xi A_\Xi'}{A_\Xi'} - \frac{1}{1+\Xi},$$

$$= \Xi\partial_\Xi \log A_\Xi' - \frac{1}{1+\Xi},$$

$$= 1 - S_{A'} - 1 + \Xi - \Xi^2 + \Xi^3 - \dots,$$

$$= -S_{A'} + \partial_{A_0'} + \partial_{A_1'} + \partial_{A_2'} + \dots;$$

a result which shews that

$$R_{a'} - (\partial_{a_0'} + \partial_{a_1'} + \partial_{a_2'} + \dots) \text{ is an invariant of even order,}$$

and $$2S_{a'} - R_{a'} - (\partial_{a_0'} + \partial_{a_1'} + \partial_{a_2'} + \dots) \text{ an invariant of uneven order.}$$

Moreover, since $$2\partial_{a_0'} + \partial_{a_1'} + \partial_{a_2'} + \dots,$$
$$\partial_{a_1'} + \partial_{a_2'} + \dots.$$

are invariants of even and uneven order respectively, we find invariants

$$R_{a'} + \partial_{a_0'} \text{ of even order,}$$

$$2S_{a'} - R_{a'} - \partial_{a_0'} \text{ of uneven order.}$$

Equivalent to these we find invariants

$$x^{-\frac{3}{2}}(R_a + \partial_{a_0}) \text{ of even order,}$$

$$x^{-\frac{1}{2}}(2S_a + x^{-1}R_a + x^{-1}\partial_{a_0}) \text{ of uneven order.}$$

The interest of these results lies in the circumstance that $S_a$ is what Sylvester has termed a Reversor in the theory of pure reciprocants. It was discovered by Hammond. The allied operator $R_a$ appears here for the first time.

If we apply the operator $x^{-\frac{3}{2}}(R_a + \partial_{a_0})$ of even order to the invariant

$$2x^{\frac{5}{2}}a_1 + x^{\frac{3}{2}}a_0 \text{ of uneven order,}$$

we obtain an invariant of uneven order. The verification is

$$x^{-\frac{3}{2}}(R_a+\partial_{a_0})(2x^{\frac{5}{2}}a_1+x^{\frac{3}{2}}a_0)=\frac{2xa_1+a_0}{a_0}=-\frac{2XA_1+A_0}{A_0}.$$

*The Invariant Reversors*

$$a_1'\partial_{a_0'}+a_2'\partial_{a_1'}+a_3'\partial_{a_2'}+\ldots=p_{a'}=-xp_a,$$
$$a_2'\partial_{a_1'}+2a_3'\partial_{a_2'}+\ldots=q_{a'}=-xq_a.$$

18. We have
$$p_{a'}=-\frac{p\xi}{1+p\xi}=\frac{(1-P)\Xi}{1+P\Xi}=I_{A'}-p_{A'};$$
establishing that
$$2p_{a'}-I_a$$
$$\equiv-2xp_a-I_a,$$
is an invariant of uneven order.

Also
$$q_{a'}=\frac{p^2\xi^2}{(1+p\xi)^2}=(1-P)^2\frac{\Xi^2}{(1+P\Xi)^2},$$
$$=(A_0'-2A_1'+A_2')\Xi^2-2(A_1'-2A_2'+A_0')\Xi^3+3(A_2'-2A_3'+A_4')\Xi^4-\ldots,$$
$$=-q_{A'}+2W_{A'}-\Omega_{A'};$$
shewing that
$$q_{a'}-W_{a'}\equiv-xq_a-W_a$$
is an invariant of uneven order.

$p_a$ and $q_a$ are therefore generating operators for the transformation, the former when the operand is homogeneous, the latter when it is isobaric.

If $j$ be the weight of the highest letter in the operand, the latter is a full invariant if
$$jI_a-2W_a$$
causes the operand to vanish.

Also
$$jp_{a'}-q_{a'}=-(jp_{A'}-q_{A'})+(jI_{a'}-2W_{a'})+\Omega_{A'},$$
shewing that
$$2(jp_{a'}-q_{a'})-(jI_{a'}-2W_{a'})$$
$$\equiv-2x(jp_a-q_a)-(jI_a-2W_a)$$
is an invariant of uneven order. Hence when the operand is a full invariant
$$+x(jp_a-q_a)$$
generates an invariant of uneven order.

*The Invariant Generators*

$$P_{a'}=(a_0'a_2'-a_1'^2)\partial_{a_1'}+(a_0'a_3'-a_1'a_2')\partial_{a_2'}+\ldots,$$
$$Q_{a'}=(a_0'a_2'-2a_1'^2)\partial_{a_1'}+2(a_0'a_3'-2a_1'a_2')\partial_{a_2'}+3(a_0'a_2'-2a_1'a_3')\partial_{a_3'}+\ldots.$$

19. These operators generate invariants in the theory of the binary quantic.

Here
$$P_{a'}=a_0'p_{a'}-a_1'I_a;$$
whence
$$P_{a'}=A_0'(I_{A'}-p_{A'})-(A_0'-A_1')I_{A'}=-A_0'p_{A'}+A_1'I_{A'}=-P_{A'};$$
shewing that
$$P_a=-x^{\frac{5}{2}}P_a$$
is an invariant of uneven order for the Halphenian substitution.

Also $$Q_{a'} = a_0' q_{a'} - 2a_1' W_{a'} = - A_0' q_{A'} + 2A_1' W_{A'} + (A_0' - 2A_1')\,\Omega_{A'},$$

or $$Q_{a'} = - Q_{A'} + (A_0' - 2A_1')\,\Omega_{A'};$$

establishing, since $a_0'\partial_{a_1'}$ is an invariant of uneven order, that

$$Q_{a'} + a_1'\Omega_{a'} \equiv - x^{\frac{3}{2}}(xQ_a - a_1\Omega_a)$$

is an invariant of uneven order.

*The Pure Reciprocant Generator*

$$G_{a'} = 4(a_0'a_2' - a_1'^2)\,\partial_{a_1'} + 5(a_0'a_3' - a_1'a_2')\,\partial_{a_2'} + 6(a_0'a_4' - a_1'a_3')\,\partial_{a_3'} + \ldots.$$

20. We find that

$$G_{a'} = Q_{a'} + 3P_{a'} + a_1' W_{a'} = - G_{A'} + A_0' W_{A'} - A_1'\Omega_{A'}.$$

Hence the invariant of uneven order

$$2G_{a'} - a_0' W_{a'} + a_1'\Omega_{a'}.$$

To verify this result take as operand

$$a_0'a_2' - a_1'^2.$$

We have
$$2G_{a'}(a_0'a_2' - a_1'^2) = 10(a_0'^2a_3' - 3a_0'a_1'a_2' + 2a_1'^3) + 4a_1'(a_0'a_2' - a_1'^2),$$
$$a_0' W_{a'}(a_0'a_2' - a_1'^2) = 2a_0'(a_0'a_2' - a_1'^2),$$
$$\Omega_{a'}(a_0'a_2' - a_1'^2) = 0.$$

Hence the result of the operation is

$$10(a_0'^2a_3' - 3a_0'a_1'a_2' + 2a_1'^3) - 2(a_0' - 2a_1')(a_0'a_2' - a_1'^2),$$

which is an invariant of uneven order because

$$a_0'^2a_3' - 3a_0'a_1'a_2' + 2a_1'^3 \text{ and } a_0' - 2a_1' \text{ are so,}$$

and $a_0'a_2' - a_1'^2$ is an invariant of even order.

$G_a$ by itself is a generator also in the theory of seminvariants when the operand is a combination of seminvariants of weight zero; for then

$$2G_{a'} - a_0' W_{a'} + a_1'\Omega_{a'}$$

is equivalent to $2G_{a'}$.

Thus the reader may verify that

$$G_a \frac{a_0^2a_3 - 3a_0a_1a_2 + 2a_1^3}{(a_0a_2 - a_1^2)^{\frac{3}{2}}}$$

is an invariant of even order; for the operand is a seminvariant of uneven order and of weight zero.

## SECTION II.

### PARTICULAR CASE OF THE $b$ TRANSFORMATION.

21. Write as usual

$$\frac{1}{m}(a_0 + a_1 u + a_2 u^2 + \ldots)^m = a_{m0} + a_{m1} u + a_{m2} u^2 + \ldots,$$

and consider the operator

$$\mu a_{m0} \partial_{a_n} + (\mu + \nu) a_{m1} \partial_{a_{n+1}} + (\mu + 2\nu) a_{m2} \partial_{a_{n+2}} + \ldots,$$

which it has been convenient to denote by

$$(\mu, \nu; m, n)_a.$$

As was shewn in a previous paper* the substitution

$$a_0 = \frac{1}{b_0},$$

$$a_1 = -\frac{b_1}{b_0^3},$$

$$a_2 = -\frac{b_2}{b_0^4} + 2\frac{b_1^2}{b_0^5},$$

$$a_3 = -\frac{b_3}{b_0^5} + 5\frac{b_1 b_2}{b_0^6} - 5\frac{b_1^3}{b_0^7},$$

..............................

which is *derived from* the formulæ for the interchange of the dependent and independent variables in the differential coefficients

$$\frac{dy}{dx}, \frac{1}{2!}\frac{d^2y}{dx^2}, \frac{1}{3!}\frac{d^3y}{dx^3}, \ldots$$

converts $(\mu, \nu; m, n)_a$

into $-\frac{1}{m}\{\mu(n-m+2), \mu - m\nu; n-m+2, n\}_b.$

An exception however occurs when $m = n+2$, for then $n - m + 2 = 0$. I refer to Art. 18, p. 156 of the paper (*loc. cit.*) where the transformation of

$$(\mu, \nu; m, n)_t$$

by the Sylvester substitution was considered. The symbolic form there given, viz.

$$\left(\frac{\mu}{m} - \nu\right) \xi^{n-m+2} \eta^m + \nu \xi^{n-m+3} \eta^{m-1} \eta',$$

becomes, when $m = n+3$,

$$\left(\frac{\mu}{n+3} - \nu\right) \frac{1}{\xi} \eta^{n+3} + \nu \eta^{n+2} \eta',$$

which may be written $\left(\frac{\mu}{n+3} - \nu\right) \partial_\eta \log(\xi) . \eta^{n+3} \eta' + \nu \eta^{n+2} \eta'.$

* "The Operator Reciprocants of Sylvester's Theory of Reciprocants," *Trans. Camb. Phil. Soc.* Vol. XXI. No. VI., 1908.

Now
$$\partial_\eta \log \xi = \partial_\eta \log (\tau\eta + \alpha_0\eta^2 + \alpha_1\eta^3 + \ldots)$$
$$= \frac{1}{\eta} + \partial_\eta \log\left(1 + \frac{\alpha_0}{\tau}\eta + \frac{\alpha_1}{\tau}\eta^2 + \ldots\right),$$
$$= \frac{1}{\eta} + \partial_\eta\left(\frac{\alpha_0}{\tau}\eta + \frac{1}{2}\frac{2\tau\alpha_1 - \alpha_0^2}{\tau^2}\eta^2 + \frac{1}{3}\frac{3\tau^2\alpha_2 - 3\tau\alpha_0\alpha_1 + \alpha_0^3}{\tau^3} + \ldots\right),$$
$$= \frac{1}{\eta} + \frac{\alpha_0}{\tau} + \frac{2\tau\alpha_1 - \alpha_0^2}{\tau^2}\eta + \frac{3\tau^2\alpha_2 - 3\tau\alpha_0\alpha_1 + \alpha_0^3}{\tau^3}\eta^2 + \ldots,$$

so that the symbolic form becomes
$$\frac{\mu}{n+3}\eta^{n+2}\eta' + \left(\frac{\mu}{n+3} - \nu\right)\left(\frac{\alpha_0}{\tau}\eta^{n+3} + \frac{2\tau\alpha_1 - \alpha_0^2}{\tau^2}\eta^{n+4} + \frac{3\tau^2\alpha_2 - 3\tau\alpha_0\alpha_1 + \alpha_0^3}{\tau^3}\eta^{n+5} + \ldots\right)\eta';$$

and now writing $-\partial_{a_{r-2}}$ for $\eta^r\eta'$ we find
$$-\frac{\mu}{n+3}\partial_{a_n} - \left(\frac{\mu}{n+3} - \nu\right)\left(\frac{\alpha_0}{\tau}\partial_{a_{n+1}} + \frac{2\tau\alpha_1 - \alpha_0^2}{\tau^2}\partial_{a_{n+2}} + \frac{3\tau^2\alpha_2 - 3\tau\alpha_0\alpha_1 + \alpha_0^3}{\tau^3}\partial_{a_{n+3}} + \ldots\right)$$
for the transform of $(\mu, \nu; n+3, n)_t$ by the Sylvester substitution.

We now make a unit increase of suffix throughout, writing
$$\alpha_0, \alpha_1, \alpha_2, \ldots \text{ for } \tau, \alpha_0, \alpha_1, \ldots,$$
then write $b$ for $\alpha$ and $n-1$ for $n$ when we find that
$$(\mu, \nu; n+2, n)_a$$
is transformed by the $b$ substitution into
$$-\frac{\mu}{n+2}\partial_{b_n} - \left(\frac{\mu}{n+2} - \nu\right)\left(\frac{b_1}{b_0}\partial_{b_{n+1}} + \frac{2b_0b_2 - b_1^2}{b_0^2}\partial_{b_{n+2}} + \frac{3b_0^2b_3 - 3b_0b_1b_2 + b_1^3}{b_0^3}\partial_{b_{n+3}} + \ldots\right)$$
and we derive the particular cases
$$(0, 1; n+2, n)_a$$
$$= \frac{b_1}{b_0}\partial_{b_{n+1}} + \frac{2b_0b_2 - b_1^2}{b_0^2}\partial_{b_{n+2}} + \frac{3b_0^2b_3 - 3b_0b_1b_2 + b_1^3}{b_0^3}\partial_{b_{n+3}} + \ldots;$$
$$(n+2, 1; n+2, n)_a$$
$$= -\partial_{b_n}.$$

*The h transformation.*

22. If
$$\frac{1}{1 - a_1u + a_2u^2 - a_3u^3 + \ldots} = 1 + h_1u + h_2u^2 + h_3u^3 + \ldots,$$
then
$$a_1 = h_1,$$
$$a_2 = h_1^2 - h_2,$$
$$a_3 = h_1^3 - 2h_1h_2 + h_3,$$
$$\ldots\ldots\ldots\ldots\ldots\ldots$$
and, putting $a_0 = 1$, I examine the effect of making this substitution upon the operator
$$(\mu, \nu; m, n)_a.$$

We have

$$\partial_{a_s}\frac{1}{1-a_1u+a_2u^2-\ldots}=\frac{(-)^{s+1}u^s}{(1-a_1u+a_2u^2-\ldots)^2}=u^s\partial_{a_s}h_s+u^{s+1}\partial_{a_s}h_{s+1}+\ldots,$$

or

$$(-)^{s+1}(1+h_1u+h_2u^2+\ldots)^2=\partial_{a_s}h_s+u\partial_{a_s}h_{s+1}+u^2\partial_{a_s}h_{s+2}+\ldots.$$

So that comparison of the coefficients of $u^p$ on either side gives us the value of $\partial_{a_s}h_{s+p}$ as a quadratic function of $h$.

Now

$$\partial_{a_s}=\frac{dh_s}{da_s}\cdot\partial_{h_s}+\frac{dh_{s+1}}{da_s}\cdot\partial_{h_{s+1}}+\frac{dh_{s+2}}{da_s}\cdot\partial_{h_{s+2}}+\ldots,$$

and we find $\partial_{a_s}=(-)^{s+1}\{\partial_{h_s}+2h_1\partial_{h_{s+1}}+(h_1^2+2h_2)\partial_{h_{s+2}}+(2h_1h_2+2h_3)\partial_{h_{s+3}}+\ldots\}.$

If we write symbolically $\partial_{h_s}=k^s,$

$$\partial_{a_s}=(-)^{s+1}k^s(1+h_1k+h_2k^2+h_3k^3+\ldots)^2,$$

and

$$a_{m0}\partial_{a_n}+a_{m1}\partial_{a_{n+1}}+a_{m2}\partial_{a_{n+2}}+\ldots$$

$$=(-)^{n+1}k^n(1+h_1k+h_2k^2+h_3k^3+\ldots)^2(a_{m0}-a_{m1}k+a_{m2}k^2-a_{m3}k^3+\ldots);$$

but

$$a_{m0}-a_{m1}k+a_{m2}k^2-\ldots=\frac{1}{m}(1+h_1k+h_2k^2+h_3k^3+\ldots)^{-m},$$

giving us

$$a_{m0}\partial_{a_n}+a_{m1}\partial_{a_{n+1}}+a_{m2}\partial_{a_{n+2}}+\ldots$$

$$=(-)^{n+1}\frac{1}{m}k^n(1+h_1k+h_2k^2+h_3k^3+\ldots)^{2-m},$$

$$=(-)^{n+1}\frac{2-m}{m}k^n(h_{2-m,0}+h_{2-m,1}k+h_{2-m,2}k^2+h_{2-m,3}k^3+\ldots),$$

leading to the relation

$$(1,0;\ m,n)_a=(-)^{n+1}\frac{2-m}{m}(1,0;\ 2-m,n)_h.$$

Again

$$(0,1;\ m,n)_a$$

$$=a_{m1}\partial_{a_{n+1}}+2a_{m2}\partial_{a_{n+2}}+3a_{m3}\partial_{a_{n+3}}+\ldots$$

$$=(-)^n k^{n+1}(1+h_1k+h_2k^2+h_3k^3+\ldots)^2(a_{m1}-2a_{m2}k+3a_{m3}k^2-\ldots);$$

but

$$a_{m0}-a_{m1}k+a_{m2}k^2-\ldots=\frac{1}{m}(1+h_1k+h_2k^2+\ldots)^{-m},$$

so that by differentiation

$$a_{m1}-2a_{m2}k+3a_{m3}k^2-\ldots=-\frac{1}{m}\partial_k(1+h_1k+h_2k^2+\ldots)^{-m},$$

and

$$(1+h_1k+h_2k^2+\ldots)^2(a_{m1}-2a_{m2}k+3a_{m3}k^2-\ldots)=\frac{1}{2-m}\partial_k(1+h_1k+h_2k^2+\ldots)^{2-m}$$

$$=h_{2-m,1}+2h_{2-m,2}k+3h_{2-m,3}k^2+\ldots,$$

and we now gather that

$$(0,1;\ m,n)_a=(-)^n(0,1;\ 2-m,n)_h;$$

and, since

$$(\mu,\nu;\ m,n)_a=\mu(1,0;\ m,n)_a+\nu(0,1;\ m,n)_a,$$

$$(\mu,\nu;\ m,n)_a=(-)^n\left\{\frac{\mu}{m}(m-2),\nu;\ 2-m,n\right\}_h,$$

it being understood that the value $m=2$ is excluded.

23. The particular case $m=2$ requires a separate examination. Reference to the foregoing investigation shews that

$$(1, 0;\ 2, n)_a$$

is symbolically

$$(-)^{n+1}\tfrac{1}{2}k^n,$$

so that

$$(1, 0;\ 2, n)_a = (-)^{n+1}\tfrac{1}{2}\partial_{h_n},$$

whilst

$$(0, 1;\ 2, n)_a = (-)^n k^{n+1}\partial_k \log(1 + h_1 k + h_2 k^2 + \dots)$$

$$= (-)^n k^{n+1}\{h_1 + (2h_2 - h_1^2)k + (3h_3 - 3h_1h_2 + h_1^3)k^2 + \dots\}$$

$$= (-)^n \{h_1 \partial_{h_{n+1}} + (2h_2 - h_1^2)\partial_{h_{n+2}} + (3h_3 - 3h_1h_2 + h_1^3)\partial_{h_{n+3}} + \dots\}.$$

Hence

$$(\mu, \nu;\ 2, n)_a$$

$$= (-)^{n+1}\frac{\mu}{m}\partial_{h_n} + \nu(-)^n\{h_1\partial_{h_{n+1}} + (2h_2 - h_1^2)\partial_{h_{n+2}} + (3h_3 - 3h_1h_2 + h_1^3)\partial_{h_{n+3}} + \dots\}.$$

*Particular Cases of the h transformation.*

24. The seminvariant annihilator

$$(1, 1;\ 1, 1)_a$$

becomes

$$(1, -1;\ 1, 1)_h,$$

or

$$\partial_{h_1} - h_2\partial_{h_3} - 2h_3\partial_{h_4} - 3h_4\partial_{h_5} - \dots.$$

Hence any seminvariant *quâ* the elements

$$h_2, h_3, h_4, \dots$$

is a seminvariant *quâ* the elements

$$a_0, a_1, a_2, \dots.$$

Ex. gr. $h_2h_4 - h_3^2$ is a seminvariant.

Moreover, writing

$$\phi_s(h_2, h_3, h_4, \dots) = \phi_s,$$

if the general solution of

$$\partial_{h_1} - h_2\partial_{h_3} - 2h_3\partial_{h_4} - 3h_4\partial_{h_5} - \dots = 0$$

be written

$$h_1^s\phi_s + \binom{s}{1}h_1^{s-1}\phi_{s-1} + \binom{s}{2}h_1^{s-2}\phi_{s-2} + \dots,$$

we have

$$sh_1^{s-1}\phi_s + s(s-1)h_1^{s-2}\phi_{s-1} + \tfrac{1}{2}s(s-1)(s-2)h_1^{s-3}\phi_{s-2} + \dots,$$

$$= h_1^s H\phi_s + \binom{s}{1}h_1^{s-1}H\phi_{s-1} + \binom{s}{2}h_1^{s-2}H\phi_{s-2} + \dots,$$

where

$$H = h_2\partial_{h_3} + 2h_3\partial_{h_4} + 3h_4\partial_{h_5} + \dots,$$

and now equating coefficients of like powers of $h_1$,

$$H\phi_s = 0,$$

$$H\phi_{s-1} = \phi_s,$$

$$H\phi_{s-2} = 2\phi_{s-1},$$

$$H\phi_{s-3} = 3\phi_{s-2},$$

$$\dots\dots\dots\dots$$

$$H\phi_{s-t} = t\phi_{s-t+1},$$

$$\dots\dots\dots\dots$$

indicating that, regarding the general solution as a binary $s^{ic}$ in $h_1$, 1, all of its seminvariants are seminvariants in the elements

$$h_2,\ h_3,\ h_4, \ldots,$$

and therefore also in the elements $a_0$, $a_1$, $a_2$, ....

25. As another interesting particular case we find that the pure reciprocant annihilator

$$(4,\ 1;\ 2,\ 1)_a$$

$$= 2\partial_{h_1} - h_1\partial_{h_2} - (2h_2 - h_1^2)\partial_{h_3} - (3h_3 - 3h_1h_2 + h_1^3)\partial_{h_4} - \ldots$$

Moreover it will be shewn that $(4,\ 1;\ 2,\ 1)_a$ can be transformed into $(1,\ 0;\ 2,\ 1)_a$, and this becomes

$$\tfrac{1}{2}\partial_{h_1},$$

of which the fundamental solutions (it being equated to zero) are

$$h_2,\ h_3,\ h_4,\ h_5, \ldots.$$

The fact is that by means of the $b$ transformation

$$a_0 = \frac{1}{b_0},\quad a_1 = -\frac{b_1}{b_0^3},\quad a_2 = -\frac{b_2}{b_0^4} + 2\frac{b_1^2}{b_0^5} \ldots$$

$(4,\ 1;\ 2,\ 1)_a$ is transformed to $-(2,\ 1;\ 1,\ 1)_b$,

and if we write $b_s = (s+1)c_s$, it is further transformed to

$$(1,\ 1;\ 1,\ 1)_c.$$

We now again employ the $b$ transformation

$$c_0 = \frac{1}{d_0},\quad c_1 = -\frac{d_1}{d_0^3},\quad c_2 = -\frac{d_2}{d_0^4} + 2\frac{d_1^2}{d_0^5}, \ldots$$

when it becomes

$$2(1,\ 0;\ 2,\ 1)_d.$$

The $h$ transformation

$$d_0 = e_0,\quad d_1 = e_1,\quad d_2 = e_1^2 - e_0e_2,\quad d_3 = e_1^3 - 2e_0e_1e_2 + e_3, \ldots$$

finally converts it into

$$e_0\partial_{e_1},$$

and we have the complete set of solutions

$$e_0,\ e_2,\ e_3,\ e_4, \ldots.$$

In detail,

$$a_0 = \frac{1}{b_0} = \frac{1}{c_0} = d_0 = e_0,$$

$$a_1 = -\frac{b_1}{b_0^3} = -\frac{2c_1}{c_0^3} = 2d_1 = 2e_1,$$

$$a_2 = -\frac{b_2}{b_0^4} + 2\frac{b_1^2}{b_0^5} = -3\frac{c_2}{c_0^4} + 8\frac{c_1}{c_0}$$

$$= 3d_2 + 2\frac{d_1^2}{d_0} = \frac{1}{e_0}(5e_1^2 - 3e_0e_2).$$

$$a_3 = -\frac{b_3}{b_0^5} + 5\frac{b_1 b_2}{b_0^6} - 5\frac{b_1^3}{b_0^7} = -4\frac{c_3}{c_0^5} + 30\frac{c_1 c_2}{c_0^6} - 40\frac{c_1^3}{c_0^7}$$

$$= 4d_3 + 10\frac{d_1 d_2}{d_0} = \frac{2}{e_0^2}(7e_1^3 - 9e_0 e_1 e_2 + 2e_0^2 e_3),$$

$$a_4 = -\frac{b_4}{b_0^6} + 6\frac{b_1 b_3}{b_0^7} + 3\frac{b_2^2}{b_0^7} - 21\frac{b_1^2 b_2}{b_0^8} + 14\frac{b_1^4}{b_0^9}$$

$$= -5\frac{c_4}{c_0^6} + 48\frac{c_1 c_3}{c_0^7} + 27\frac{c_2^2}{c_0^7} - 252\frac{c_1^2 c_2}{c_0^8} + 224\frac{c_1^4}{c_0^9}$$

$$= 5d_4 + 18\frac{d_1 d_3}{d_0} + 12\frac{d_2^2}{d_0} + 9\frac{d_1^2 d_2}{d_0^2} - 2\frac{d_1^4}{d_0^3}$$

$$= \frac{1}{e_0^3}(42e_1^4 - 84e_0 e_1^2 e_2 + 17e_0^2 e_2^2 + 28e_0^2 e_1 e_3 - 5e_0^3 e_4),$$

&c.

This then is the transformation from the elements $a$ to the elements $e$ which transforms the operator. As above shewn,

$$e_0,\ e_2,\ e_3,\ e_4,\ \ldots$$

are pure reciprocants, and we find by calculation their expressions in terms of the elements $a$. Viz.

$$e_0 = a_0,$$

$$e_2 = \frac{1}{12a_0}(5a_1^2 - 4a_0 a_2),$$

$$e_3 = \frac{1}{4a_0^2}(a_0^2 a_3 - 3a_0 a_1 a_2 + 2a_1^3),$$

$$e_4 = \frac{1}{720a_0^3}(551a_1^4 - 1184a_0 a_1^2 a_2 + 272a_0^2 a_2^2 + 504a_0^2 a_1 a_3 - 144a_0^3 a_4),$$

&c.

The first three of these will be immediately recognized.

The last one, in the bracket, is expressed in terms of Sylvester's ground forms by the formula

$$\tfrac{551}{25}(5a_1^2 - 4a_0 a_2)^2 - \tfrac{72}{25}(105a_0 a_1^2 a_2 + 28a_0^2 a_2^2 - 175a_0^2 a_1 a_3 + 50a_0^3 a_4).$$

The forms obtained in the above manner form a complete set from which all pure reciprocants can be obtained, but they do not constitute the simplest set of ground forms obtainable.

The possibility of such a transformation depends upon two circumstances. In the first place the operator

$$(4,\ 1;\ 2,\ 1)$$

is a particular case of the operator

$$(\mu,\ \nu;\ m,\ n)$$

such that $n - m + 2 = 1$; consequently the $b$ transformation results in an operator, viz.

$$(2,\ 1;\ 1,\ 1),$$

which is linear in the elements which are coefficients of the differential inverses. In the second place any such linear operator $(\mu, \nu; 1, n)$ can be transformed by mere numerical multiplication of the elements into any other form $(\mu', \nu'; 1, n)$. It follows that $(2, 1; 1, 1)$ is transformable into the seminvariant operator $(1, 1; 1, 1)$. An immediate consequence of this, of course, is that all pure reciprocants are transformable into seminvariants. The $b$ transformation now transforms $(1, 1; 1, 1)$ into an operator, viz. $(1, 0; 2, 1)$, in which the element $\nu$ is zero, and this being so the $h$ transformation produces the final simple form of operator

$$e_0 \partial_{e_1}.$$

*A transformation of the Seminvariant Operator.*

26. It will be observed that the seminvariant operator $(1, 1; 1, 1)$ is by the successive $b$ and $h$ substitutions brought to the required simple form. This in fact takes place during the transformation of the pure reciprocant operator.

The transformation which effects this is

$$a_1 = -c_1,$$
$$a_2 = c_1^2 + c_2,$$
$$a_3 = -c_1^3 - 3c_1c_2 - c_3,$$
$$a_4 = c_1^4 + 6c_1^2c_2 + 2c_2^2 + 4c_1c_3 + c_4,$$
&c.

In general $$a_n = (-)^n \Sigma \frac{n!}{(\lambda_0 + 1)!\,\lambda_1!\,\lambda_2!\ldots} c_0^{\lambda_0} c_1^{\lambda_1} c_2^{\lambda_2} \ldots.$$

The operator then becomes $$c_0 \partial_{c_1},$$
and $c_0, c_2, c_3, c_4, \ldots$ are seminvariants.

In fact from the above relations

$$c_2 = -a_1^2 + a_2,$$
$$c_3 = -2a_1^3 + 3a_1a_2 - a_3,$$
$$c_4 = -5(-a_1^2 + a_2)^2 + 3a_2^2 - 4a_1a_3 + a_4,$$
..............................................

and in general $$c_n = \Sigma (-)^{n-\Sigma\lambda+1} \frac{(n + \Sigma\lambda - 2)!}{(n-1)!\,\lambda_1!\,\lambda_2!\ldots} a_1^{\lambda_1} a_2^{\lambda_2} \ldots.$$

In the expressions of $c_2, c_3, c_4$, seminvariants are at once recognizable.

*The s transformation.*

27. If $s_1, s_2, s_3, \ldots$ denote the sums of powers of the roots of the equation

$$x^n - a_1x^{n-1} + a_2x^{n-2} - \ldots = 0, \text{ where } n = \infty,$$

it will be shewn that

$$V_a = (4, 1; 2, 1)_a = 2(\partial_{s_1} - 3a_1\partial_{s_2} + 6a_2\partial_{s_3} - 10a_3\partial_{s_4} - \ldots),$$
$$\Omega_a = (1, 1; 1, 1)_a = \partial_{s_1} - 2s_1\partial_{s_2} - 3s_2\partial_{s_3} - 4s_3\partial_{s_4} - \ldots.$$

To prove the first of these relations, it is easy to shew that

$$\partial_{a_p} = (-)^{p+1}\{p\partial_{s_p} + (p+1)h_1\partial_{s_{p+1}} + (p+2)h_2\partial_{s_{p+2}} + \ldots\},$$

leading to

$$\begin{aligned} 2V_a = & \, 4a_{20}(\partial_{s_1} + 2h_1\partial_{s_2} + 3h_2\partial_{s_3} + 4h_3\partial_{s_4} + \ldots) \\ & - 5a_{21}(2\partial_{s_2} + 3h_1\partial_{s_3} + 4h_2\partial_{s_4} + \ldots) \\ & + 6a_{22}(3\partial_{s_3} + 4h_1\partial_{s_4} + 5h_2\partial_{s_5} + \ldots) \\ & - 7a_{23}(4\partial_{s_4} + 5h_1\partial_{s_5} + \ldots) \\ & + \ldots . \end{aligned}$$

Herein the coefficient of $\partial_{s_p}$ is

$$p\{4a_{20}h_{p-1} - 5a_{21}h_{p-2} + 6a_{22}h_{p-3} - \ldots + (-)^{p-1}(p+3)a_{2,p-1}\}.$$

Now if $$\frac{1}{1 - a_1u + a_2u^2 - \ldots} = 1 + h_1u + h_2u^2 + \ldots,$$

which we may write in the form

$$A_u^{-1} = H_u,$$

$$4A_u^2 + u\partial_u A_u^2 = 4a_{20} - 5a_{21}u + 6a_{22}u^2 - \ldots,$$

and $$A_u^{-1} = 1 + h_1u + h_2u^2 + \ldots;$$

therefore by multiplication

$$\begin{aligned} & 4a_{20} + (4a_{20}h_1 - 5a_{21})u + (4a_{20}h_2 - 5a_{21}h_1 + 6a_{22})u^2 - \ldots \\ & \quad + \{4a_{20}h_{p-1} - 5a_{21}h_{p-2} + \ldots + (-)^{p+1}(p+3)a_{2,p-1}\}u^{p-1} + \ldots, \\ & \quad = 4A_u + 2u\partial_u A_u, \\ & \quad = 4(1 - a_1u + a_2u^2 - a_3u^3 + \ldots) + (-2a_1u + 4a_2u^2 - 6a_3u^3 + \ldots), \\ & \quad = 4 - 6a_1u + 8a_2u^2 - 10a_3u^3 + \ldots . \end{aligned}$$

In this series the coefficient of $u^{p-1}$ is

$$(-)^{p+1}(2p+2)a_{p-1},$$

proving that the coefficient of $\partial_{s_p}$ is

$$(-)^{p+1}p(2p+2)a_{p-1}.$$

Therefore $$2V_a = \Sigma(-)^{p-1}p(2p+2)a_{p-1}\partial_{s_p},$$

or $$V_a = 2\Sigma(-)^{p+1}\binom{p+1}{2}a_{p-1}\partial_{s_p},$$

or $$V_a = 2(\partial_{s_1} - 3a_1\partial_{s_2} + 6a_2\partial_{s_3} - 10a_3\partial_{s_4} + \ldots).$$

For the second relation

$$\Omega_a = \partial_{s_1} + 2(h_1 - 2a_1)\partial_{s_2} + 3(h_2 - 2a_1h_1 + 3a_2)\partial_{s_3} + 4(h_3 - 2a_1h_2 + 3a_2h_1 - 4a_3)\partial_{s_4} + \ldots,$$

and by easy algebra this reduces to

$$\Omega_a = \partial_{s_1} - 2s_1\partial_{s_2} - 3s_2\partial_{s_3} - 4s_3\partial_{s_4} - \ldots .$$

If, however, we write $$s_p = -s_p',$$

$$\Omega_a = -\partial_{s_1'} - 2s_1'\partial_{s_2'} - 3s_2'\partial_{s_3'} - 4s_3'\partial_{s_4'} - \ldots .$$

indicating that every seminvariant in the elements

$$a_1, a_2, a_3, \ldots$$

is also a seminvariant in the elements

$$-s_1, -s_2, -s_3, \ldots.$$

Thus for example from the seminvariant

$$a_0a_4 - 4a_1a_3 + 3a_2^2$$

we at once derive a new seminvariant

$$-s_4 - 4s_1s_3 + 3s_2^2,$$

which is found to be equal to

$$-2(a_1^4 - 2a_1^2a_2 - 5a_2^2 + 8a_1a_3 - 2a_4).$$

It should be noted that the important operator

$$J_a = (2, 1; 2, 2)_a$$

is equivalent to $$-2(\partial_{s_2} - 3a_1\partial_{s_3} + 6a_2\partial_{s_4} - 10a_3\partial_{s_5} + \ldots).$$

*The combined b and h transformations.*

28. Both the $b$ and the $h$ substitutions being of period 2, we can combine the substitutions alternately. Thus if we first employ the $b$ substitution and then the $h$ and $b$ substitutions alternately, we obtain substitutions which we may denote by

$$(hb)^p \text{ and } b(hb)^{p-1}.*$$

It will be found that

$$(hb)^p(\mu, \nu; m, n) = (-)^{np}\left\{-\frac{\mu}{m}(pn - m), \frac{m\nu - p\mu}{m}; m - pn, n\right\},$$

$$b(hb)^{p-1}(\mu, \nu; m, n) = (-)^{n(p+1)}\left\{-\frac{\mu}{m}(pn - m + 2), \frac{m\nu - p\mu}{m}; pn - m + 2, n\right\}.$$

Also

$$(bh)^p(\mu, \nu; m, n) = (-)^{np}\left\{\frac{\mu}{m}(pn + m), \frac{m\nu + p\mu}{m}; m + pn, n\right\},$$

$$h(bh)^p(\mu, \nu; m, n) = (-)^{n(p+1)}\left\{\frac{\mu}{m}(pn + m - 2), \frac{m\nu + p\mu}{m}; 2 - m - pn, n\right\}.$$

In the first, third and fourth of these formulæ $p$ may be zero or any positive integer; in the second $p$ may be any positive integer, zero excluded. Observing that the third and fourth operators become respectively equal to the first and second when $-p$ is written for $p$, we gather that we only require the third and fourth operators in which $p$ may be supposed to be zero or any positive or negative integer.

As a particular case we find that the seminvariant operator

$$(1, 1; 1, 1)$$

may be transformed into $$(-)^p(p + 1, p + 1; p + 1, 1),$$

and into $$(-)^{p+1}(p - 1, p + 1; 1 - p, 1);$$

* The substitutions are from right to left successively.

and the pure reciprocant operator (4, 1 ; 2, 1) into

$$(-)^p\,(2p+4,\ 2p+1\ ;\ p+2,\ 1),$$

or

$$(-)^{p+1}\,(2p,\ 2p+1\ ;\ -p,\ 1).$$

*Transformation by suffix diminution.*

29. The operator $(\mu,\ \nu\ ;\ m,\ n)$ admits of one very simple transformation which may be repeated indefinitely. If therein we put

$$a_0 = 0,\quad a_s = a_{s-1}$$

for all values of $s$, or in other words if we diminish each suffix by unity, the operator becomes

$$(\mu + m\nu,\ \nu\ ;\ m,\ m+n-1),$$

and the solutions of this operator are obtained from those of $(\mu,\ \nu\ ;\ m,\ n)$ by subjecting the solutions to a unit diminution of suffix.

If we employ this transformation $p$ times we reach the operator

$$(\mu + pm\nu,\ \nu\ ;\ m,\ pm+n-p).$$

This operator is effectively the same as the untransformed operator when

$$m = 1,\ \ \nu = 0.$$

# VIII. *Vector Integral Equations and Gibbs' Dyadics.*

By C. E. Weatherburn (Ormond College, Melbourne).

[Communicated by Mr G. H. Hardy.]

[*Received* 4 October 1915—*Read* 25 October 1915.]

CONTENTS.



## Introduction.

The integral equations that have hitherto been considered by mathematicians are scalar equations in which the functions involved are not related to any particular direction in space. In the problems of mathematical physics we are frequently led by Cartesian analysis to a system of three integral equations with the same number of unknown functions, while if our methods are those of vector analysis we find instead a single integral equation in which the unknown function is a vector quantity, and the kernel of the equation no longer a scalar but either a vector used in cross multiplication or an operator involving vectors. The vector integral equations that most commonly appear are of the form

$$\mathbf{u}(t) - \lambda \int \mathbf{K}(ts) \cdot \mathbf{u}(s)\, ds = \mathbf{f}(t) \qquad (1),$$

where $\mathbf{u}(t)$ is the unknown vector function of the position of the point $t$, $\lambda$ an arbitrary parameter, $\mathbf{f}(t)$ a known vector function, and $\mathbf{K}(ts)$ a dyadic*, i.e. an operator which acting on the vector $\mathbf{u}$ gives a linear vector function of $\mathbf{u}$. It will be shewn in the following that (1) is the most general type of linear vector integral equation of the

* Cf. Gibbs-Wilson, *Vector Analysis*, New York, 1901, Chapter v.

second kind, and it is to equations of this sort, analogous to Fredholm's equation, that we shall confine our attention in the present paper.

The scalar kernel of the ordinary integral equation is thus replaced by the dyadic operator $\mathbf{K}(ts)$ which involves vector functions of the positions of the two points $t$ and $s$. In developing the theory of the equation (1) along lines suggested by the theory of Fredholm's equation some new ideas will be necessary in view of certain important differences between the algebra of dyadics and that of ordinary scalar multipliers. For instance, the fact that the commutative law does not in general hold for the factors of a dyadic product makes it essential, if we wish to introduce determinants of dyadic elements, to formulate rules for the expansion of such determinants according to which the elements in each term will occur in some definite order. This the author believes is done with complete success by the introduction of two kinds of dyadic determinants, called respectively *row* and *column determinants*, by means of which two series are formed whose quotient is the resolvent dyadic for the integral equation (1).

A special class of kernel will be considered (for which the author suggests the term *conjugo-symmetric*) which need not be either self-conjugate or symmetric but which suggests a blending of these two kinds. Such a kernel makes the vector integral equation identical with its associated equation. A set of theorems will be established for the conjugo-symmetric kernel analogous to those that hold for the symmetric kernel of the scalar integral equation.

## I. THE LINEAR VECTOR INTEGRAL EQUATION OF THE SECOND KIND.

§ 1. We are familiar with a system of three linear scalar integral equations of the second kind in the following form to which the system may always be reduced

$$\left.\begin{aligned} u_1(t) - \lambda \int [K_{11}(ts)\, u_1(s) + K_{12}(ts)\, u_2(s) + K_{13}(ts)\, u_3(s)]\, ds = f_1(t) \\ u_2(t) - \lambda \int [K_{21}(ts)\, u_1(s) + K_{22}(ts)\, u_2(s) + K_{23}(ts)\, u_3(s)]\, ds = f_2(t) \\ u_3(t) - \lambda \int [K_{31}(ts)\, u_1(s) + K_{32}(ts)\, u_2(s) + K_{33}(ts)\, u_3(s)]\, ds = f_3(t) \end{aligned}\right\} \quad \ldots\ldots\ldots\ldots\ldots (2),$$

where all the functions are scalar functions of the positions of the points indicated, $u_i(t)$ $(i = 1, 2, 3)$ are the unknowns, and the integration is to be extended over a definite fixed region $S$ which may be a line, surface, or volume, $ds$ being the element of that region surrounding the point $s$. We shall assume that the functions $f_n(t)$, $K_{nm}(ts)$ are finite within the region considered.

Let $\mathbf{i}$, $\mathbf{j}$, $\mathbf{k}$ be the unit vectors of a rectangular system. The scalar functions of (2) grouped in trios may be regarded as the tensors of the components in these directions of vectors defined by the relations

$$\mathbf{u}(t) = \mathbf{i}\, u_1(t) + \mathbf{j}\, u_2(t) + \mathbf{k}\, u_3(t),$$

$$\mathbf{f}(t) = \mathbf{i}\, f_1(t) + \mathbf{j}\, f_2(t) + \mathbf{k}\, f_3(t),$$

$$\mathbf{K}_r(ts) = \mathbf{i}\, K_{r1}(ts) + \mathbf{j}\, K_{r2}(ts) + \mathbf{k}\, K_{r3}(ts),$$

$$r = 1, 2, 3.$$

The integrands in (2) are then equal to the scalar products $\mathbf{K}_1(ts)\cdot\mathbf{u}(s)$, $\mathbf{K}_2(ts)\cdot\mathbf{u}(s)$ and $\mathbf{K}_3(ts)\cdot\mathbf{u}(s)$ respectively. On multiplying the equations (2) by $\mathbf{i}$, $\mathbf{j}$, $\mathbf{k}$ respectively and adding we have the single integral equation

$$\mathbf{u}(t)-\lambda\int[\mathbf{i}\,\mathbf{K}_1(ts)+\mathbf{j}\,\mathbf{K}_2(ts)+\mathbf{k}\,\mathbf{K}_3(ts)]\cdot\mathbf{u}(s)\,ds=\mathbf{f}(t)\ \ldots\ldots\ldots\ldots\ldots\ldots(1'),$$

which may be written

$$\mathbf{u}(t)-\lambda\int\mathbf{K}(ts)\cdot\mathbf{u}(s)\,ds=\mathbf{f}(t)\ \ldots\ldots\ldots\ldots\ldots\ldots(1),$$

where the function appearing as the kernel of the vector integral equation is the dyadic

$$\mathbf{K}(ts)=\mathbf{i}\,\mathbf{K}_1(ts)+\mathbf{j}\,\mathbf{K}_2(ts)+\mathbf{k}\,\mathbf{K}_3(ts).$$

This is an operator each term of which is the indeterminate product of two vectors known respectively as the antecedent and the consequent of that dyad. The dyadic in (1) occurs as a *prefactor* to the vector $\mathbf{u}(s)$, and the result of its operation is the sum of the products of each antecedent by the scalar product of its consequent and the vector $\mathbf{u}(s)$. When expressed in nonion form* the kernel $\mathbf{K}(ts)$ becomes

$$\begin{aligned}&K_{11}(ts)\,\mathbf{ii}+K_{12}(ts)\,\mathbf{ij}+K_{13}(ts)\,\mathbf{ik}\\+&K_{21}(ts)\,\mathbf{ji}+K_{22}(ts)\,\mathbf{jj}+K_{23}(ts)\,\mathbf{jk}\\+&K_{31}(ts)\,\mathbf{ki}+K_{32}(ts)\,\mathbf{kj}+K_{33}(ts)\,\mathbf{kk},\end{aligned}$$

the "determinant" of which† is identical with the determinant of the coefficients under the integral sign in the system (2).

§ 2. The system (2) of scalar integral equations is then equivalent to the single vector equation (1). Conversely (1) may be replaced by the system (2). We shall need frequently to refer to another integral equation intimately related to (1), viz. the vector equation

$$\mathbf{v}(t)-\lambda\int\mathbf{v}(s)\cdot\mathbf{K}(st)\,ds=\mathbf{f}(t)\ \ldots\ldots\ldots\ldots\ldots\ldots(3),$$

which will be called the *associated* equation. In this the dyadic kernel $\mathbf{K}(st)$ occurs as a *postfactor* to the vector $\mathbf{v}(s)$, so that the result of its direct operation is the sum of the products of each consequent by the scalar product of its antecedent and the vector $\mathbf{v}(s)$. When expanded the integrand in this equation becomes

$$\begin{aligned}&\mathbf{i}\,[v_1(s)K_{11}(st)+v_2(s)K_{21}(st)+v_3(s)K_{31}(st)]\\+\ &\mathbf{j}\,[v_1(s)K_{12}(st)+v_2(s)K_{22}(st)+v_3(s)K_{32}(st)]\\+\ &\mathbf{k}\,[v_1(s)K_{13}(st)+v_2(s)K_{23}(st)+v_3(s)K_{33}(st)],\end{aligned}$$

so that the vector equation (3) is equivalent to the following system of three scalar equations

$$\left.\begin{aligned}v_1(t)-\lambda\int[v_1(s)K_{11}(st)+v_2(s)K_{21}(st)+v_3(s)K_{31}(st)]\,ds=f_1(t)\\v_2(t)-\lambda\int[v_1(s)K_{12}(st)+v_2(s)K_{22}(st)+v_3(s)K_{32}(st)]\,ds=f_2(t)\\v_3(t)-\lambda\int[v_1(s)K_{13}(st)+v_2(s)K_{23}(st)+v_3(s)K_{33}(st)]\,ds=f_3(t)\end{aligned}\right\}\ \ldots\ldots\ldots(4).$$

This system, it should be observed, is not identical with (2). The rows of the coefficients in (4) agree with the columns of the coefficients in (2) with the variables $s$ and $t$ interchanged; and vice versa. It is a *common mistake* in discussing the system

* Gibbs-Wilson, *loc. cit.* p. 269. † *Ibid.* p. 317.

(2) of integral equations to assume that the associated system differs only in the interchange of variables. This assumption is quite wrong except in the special case in which $K_{ri}(ts) = K_{ir}(ts)$: that is when the dyadic $\mathbf{K}(ts)$ is self-conjugate. In this particular case it is immaterial whether the dyadic is placed as a prefactor or a postfactor. But in general the relative position of the kernel and the unknown vector cannot be varied at pleasure.

§ 3. Either of the equations (1) and (3) is the most general form of linear vector integral equation of the second kind. For it is the fundamental property of a dyadic that when operating on a vector $\mathbf{u}$ it gives a linear vector function of $\mathbf{u}$; while every linear vector function may be represented by a dyadic* to be used as a prefactor, or by the conjugate of that dyadic used as a postfactor. Hence the most general form of the integrand is the direct product of a dyadic and the unknown vector. A form that might suggest itself is $\mathbf{a} \times \mathbf{u}$, where $\mathbf{a}$ is a vector independent of $\mathbf{u}$. This form is included in the above, for a vector $\mathbf{a}$ used in cross multiplication is equivalent† to the dyadic $\mathbf{I} \times \mathbf{a}$ or $\mathbf{a} \times \mathbf{I}$ used in direct (scalar) multiplication, $\mathbf{I}$ being the idemfactor, that is the dyadic whose operation leaves a vector unchanged. The case in which $\mathbf{u}$ is multiplied by a scalar function $m$ is equivalent to that in which the dyadic is $m\mathbf{I}$.

If the unknown $\mathbf{u}$ outside the integral sign has a dyadic either as a prefactor or as a postfactor, the equation may be multiplied throughout by its reciprocal‡ dyadic and thus reduced to the form (1) or (3).

Moreover the dyadic kernel

$$\mathbf{K}(ts) = \mathbf{i}\,\mathbf{K}_1(ts) + \mathbf{j}\,\mathbf{K}_2(ts) + \mathbf{k}\,\mathbf{K}_3(ts)$$

is the most general form of dyadic. For every dyadic may be reduced to the sum of three dyads, of which either the antecedents or the consequents may be arbitrarily chosen§ provided they are not coplanar. In the present form our arbitrarily chosen antecedents are the rectangular unit vectors $\mathbf{i}$, $\mathbf{j}$, $\mathbf{k}$. We have shewn then that (1) is the most general form of linear vector integral equation of the second kind.

§ 4. In his classical memoir Fredholm‖ has shewn how a system of integral equations such as (2), in the case where the region of integration is linear, may be reduced theoretically to a single scalar integral equation whose kernel and unknown each represent different functions in various sections of the domain of integration. The possibilities of this method, such as they are, may be extended to the general case in which the region of integration, $S$, is a surface or a volume. If this be replaced by another, $S'$, consisting of the original region $S$ traversed three times the system (2) is equivalent to the single scalar integral equation

$$u(t) - \lambda \int_{S'} K(ts)\, u(s)\, ds = f(t) \quad \text{.................................(5)},$$

where, if $t$ and $s$ are points of the region being traversed for the $n$th and $m$th times respectively $(n, m = 1, 2, 3)$,

$$K(ts) = K_{nm}(ts), \quad u(t) = u_n(t), \quad u(s) = u_m(s) \ \text{ and } \ f(t) = f_n(t).$$

* Gibbs-Wilson, *loc. cit.* p. 267. † *Ibid.* p. 299. ‡ *Ibid.* p. 290. § *Ibid.* p. 271.

‖ "Sur une classe d'équations fonctionnelles," *Acta Math.* Bd. XXVII. (1903), pp. 378, 379.

From the solution $u(t)$ of this equation the unknown functions $u_1(t)$, $u_2(t)$ and $u_3(t)$ of (2) are found by the above relations.

But though this procedure reduces the system (2) theoretically to the case of a single integral equation the method is rather cumbrous in practice as all the functions involved change abruptly and frequently within the region of integration. The enquiry therefore suggests itself whether we can work with the single vector integral equation (1) to which the system has been reduced, and develop if possible the theory of vector integral equations as an important and useful branch of vector analysis. The enquiry is all the more essential to one who works habitually with vector methods, for it is in the form (1) that the integral equation presents itself to him: and it would be a doubtful gain to give up a single vector equation for a system of scalars, even if that system be reducible to a single scalar such as (5).

## II. The Iterated Dyadic Kernels, and solution by Successive Substitutions.

§ 5. Before proceeding with the solution of our equation (1) we shall introduce the idea of an iterated dyadic. The algebra of dyadics makes us familiar with the direct product of two or more dyadics. The direct (scalar) product of the dyadics $\mathbf{K}(t\vartheta)$ and $\mathbf{K}(\vartheta s)$ is written $\mathbf{K}(t\vartheta)\cdot\mathbf{K}(\vartheta s)$, and is the formal expansion of the product, according to the distributive law, as a sum of products of dyads. The product is itself a dyadic and the sum of any number of dyadics is a dyadic. Hence multiplying the product by $d\vartheta$ and summing for all the elements of the region $S$, we have in the limit that the integral

$$\int \mathbf{K}(t\vartheta)\cdot\mathbf{K}(\vartheta s)\,d\vartheta$$

is a dyadic. Further we may have the product of three or more dyadics, and the factors of such a product are known to be associative though not in general commutative. The products

$$\mathbf{K}(t\vartheta)\cdot[\mathbf{K}(\vartheta\sigma)\cdot\mathbf{K}(\sigma s)] \quad \text{and} \quad [\mathbf{K}(t\vartheta)\cdot\mathbf{K}(\vartheta\sigma)]\cdot\mathbf{K}(\sigma s)$$

are identical. Multiplying by the scalar product $d\vartheta d\sigma$ and summing over the whole region of integration for each of the variables $\vartheta$ and $\sigma$, we have in the limit

$$\int \mathbf{K}(t\vartheta)\cdot\left[\int \mathbf{K}(\vartheta\sigma)\cdot\mathbf{K}(\sigma s)\,d\sigma\right]d\vartheta = \int\left[\int \mathbf{K}(t\vartheta)\cdot\mathbf{K}(\vartheta\sigma)\,d\vartheta\right]\cdot\mathbf{K}(\sigma s)\,d\sigma \quad \ldots\ldots(6).$$

The process may be continued for any number of factors so that the order of integration may be changed at pleasure and the associative property used for any grouping of consecutive factors. The factors however are not commutative.

The dyadics formed in this way by successive iterations of the dyadic kernel $\mathbf{K}(ts)$ will be called the *iterated dyadic kernels*, or briefly the iterated kernels. The dyadic

$$\mathbf{K}_1(ts) = \int \mathbf{K}(t\vartheta)\cdot\mathbf{K}(\vartheta s)\,d\vartheta$$

will be referred to as the first iterated kernel:

$$\begin{aligned}\mathbf{K}_2(ts) &= \int \mathbf{K}_1(t\vartheta)\cdot\mathbf{K}(\vartheta s)\,d\vartheta\\ &= \int \mathbf{K}(t\vartheta)\cdot\mathbf{K}_1(\vartheta s)\,d\vartheta\end{aligned}$$

as the second, and in general

$$\begin{aligned}\mathbf{K}_p(ts) &= \int \mathbf{K}_{p-1}(t\vartheta)\cdot\mathbf{K}(\vartheta s)\,d\vartheta\\ &= \int \mathbf{K}_{p-q-1}(t\vartheta)\cdot\mathbf{K}_q(\vartheta s)\,d\vartheta\end{aligned}$$

as the $p$th iterated kernel.

The operation of the dyadic $\mathbf{K}(ts)$ on a vector $\mathbf{a}$ yields another vector $\mathbf{b}$. The ratio $|\mathbf{b}| : |\mathbf{a}|$ depends of course on the direction of the vector $\mathbf{a}$; but if the dyadic $\mathbf{K}(ts)$ is finite, that is, every element of its determinant finite, there will be a scalar function $M(ts)$ such that this modular magnification $\leqslant M(ts)$. Further for all points $t$ and $s$ of the region $S$ there will be a finite number $M$ such that $M(ts) \leqslant M$. This number $M$ we shall speak of as the *greatest modular magnification* by the dyadic $\mathbf{K}(ts)$. It follows that if $B$ is the magnitude of the region of integration, and $u_0$ the maximum value of $|\mathbf{u}(\vartheta)|$ for all points $\vartheta$ of this region

$$|\int \mathbf{K}(t\vartheta) \cdot \mathbf{u}(\vartheta)\, d\vartheta| \leqslant BMu_0,$$

and generally

$$|\int \mathbf{K}_p(t\vartheta) \cdot \mathbf{u}(\vartheta)\, d\vartheta| \leqslant (BM)^{p+1} u_0 \quad \ldots\ldots(7).$$

§ 6. We can now shew that the method of successive substitutions may, with certain restrictions on the parameter $\lambda$, be used to obtain the solution of our standard equation

$$\mathbf{u}(t) - \lambda \int \mathbf{K}(ts) \cdot \mathbf{u}(s)\, ds = \mathbf{f}(t) \quad \ldots\ldots(1).$$

For on substituting for $\mathbf{u}(s)$ under the integral sign the value given by the equation itself, we deduce

$$\mathbf{u}(t) = \mathbf{f}(t) + \lambda \int \mathbf{K}(t\vartheta) \cdot [\mathbf{f}(\vartheta) + \lambda \int \mathbf{K}(\vartheta\sigma) \cdot \mathbf{u}(\sigma)\, d\sigma]\, d\vartheta$$
$$= \mathbf{f}(t) + \lambda \int \mathbf{K}(t\vartheta) \cdot \mathbf{f}(\vartheta) + \lambda^2 \int \mathbf{K}_1(t\sigma) \cdot \mathbf{u}(\sigma)\, d\sigma.$$

Substituting in this equation the value of $\mathbf{u}(\sigma)$ given by (1) and continuing the process we find

$$\mathbf{u}(t) = \mathbf{f}(t) + \lambda \int \mathbf{K}(t\vartheta) \cdot \mathbf{f}(\vartheta)\, d\vartheta + \lambda^2 \int \mathbf{K}_1(t\vartheta) \cdot \mathbf{f}(\vartheta)\, d\vartheta + \ldots$$
$$+ \lambda^n \int \mathbf{K}_{n-1}(t\vartheta) \cdot \mathbf{f}(\vartheta)\, d\vartheta + \mathbf{R}_n \quad \ldots\ldots(8),$$

where

$$\mathbf{R}_n = \lambda^{n+1} \int \mathbf{K}_n(t\vartheta) \cdot \mathbf{u}(\vartheta)\, d\vartheta.$$

If we write the equation (8) as

$$\mathbf{u}(t) = \mathbf{S}_n(t) + \mathbf{R}_n \quad \ldots\ldots(8'),$$

it is easy to shew that, with a certain restriction on $\lambda$, the series represented by $\mathbf{S}_n(t)$ is absolutely and uniformly convergent when $n$ increases indefinitely. For if $\mathbf{S}(t)$ denote this infinite series we have in virtue of (7)

$$|\mathbf{S}(t)| \leqslant |\mathbf{f}(t)| + \sum_{n=1}^{\infty} |\lambda^n \int \mathbf{K}_{n-1}(t\vartheta) \cdot \mathbf{f}(\vartheta)\, d\vartheta|$$
$$\leqslant \sum_{n=0}^{\infty} (|\lambda| BM)^n f_0,$$

where $f_0$ is the maximum value of $|\mathbf{f}(\vartheta)|$. Now this series is convergent if $|\lambda| < \frac{1}{BM}$. If then this restriction is imposed on $\lambda$ the series $\mathbf{S}(t)$ is absolutely and uniformly convergent.

That $\mathbf{S}(t)$ actually represents the solution of (1) under these conditions will be shewn from another point of view in the following section. Here we observe that we may write

$$\mathbf{S}(t) = \mathbf{f}(t) + \lambda \int \mathbf{H}(t\vartheta) \cdot \mathbf{f}(\vartheta)\, d\vartheta \quad \ldots\ldots(9),$$

where $\mathbf{H}(t\vartheta)$ is the infinite series of dyadics

$$\mathbf{H}(t\vartheta) = \mathbf{K}(t\vartheta) + \lambda \mathbf{K}_1(t\vartheta) + \lambda^2 \mathbf{K}_2(t\vartheta) + \ldots \quad \ldots\ldots(10).$$

When $|\lambda| < \frac{1}{BM}$ this series is absolutely and uniformly convergent; that is to say, the result of its operation, term by term, on a finite vector gives an absolutely and uniformly convergent series of vector functions. This follows immediately from the above.

## III. The Resolvent Dyadic.

§ 7. The dyadic $\mathbf{H}(ts)$ defined by (10) is connected with the kernel $\mathbf{K}(ts)$ by alternative relations of a specially simple nature. For it follows immediately from (10) and the properties of the iterated kernels that

$$\mathbf{H}(ts) - \mathbf{K}(ts) = \lambda \int [\mathbf{K}(t\vartheta) + \lambda \mathbf{K}_1(t\vartheta) + \dots \text{ to } \infty] \cdot \mathbf{K}(\vartheta s)\, d\vartheta$$
$$= \lambda \int \mathbf{H}(t\vartheta) \cdot \mathbf{K}(\vartheta s)\, d\vartheta \quad \dots\dots (11),$$

and similarly that

$$\mathbf{H}(ts) - \mathbf{K}(ts) = \lambda \int \mathbf{K}(t\vartheta) \cdot [\mathbf{K}(\vartheta s) + \lambda \mathbf{K}_1(\vartheta s) + \dots \text{ to } \infty]\, d\vartheta$$
$$= \lambda \int \mathbf{K}(t\vartheta) \cdot \mathbf{H}(\vartheta s)\, d\vartheta \quad \dots\dots (11').$$

These equations shew that there is a reciprocal relation between the kernel $\mathbf{K}(ts)$ and the function $\mathbf{H}(ts)$ which we shall call the *resolvent.* But we shall avoid speaking of either as the reciprocal function of the other, because in the usual terminology two dyadics are said to be reciprocal when their product is equal to the idemfactor. The reciprocal dyadic of $\mathbf{K}(ts)$ would be denoted by $\mathbf{K}^{-1}(ts)$ and is quite different from $\mathbf{H}(ts)$.

The relations (11) and (11′) have been proved only on the assumption that $|\lambda| < \frac{1}{BM}$. But these equations have an importance for a wider range of parameter values than this. We shall prove that *if there exists a dyadic* $\mathbf{H}(ts)$ *connected with* $\mathbf{K}(ts)$ *by the relation* (11) *then the integral equation* (1) *admits a unique solution given by*

$$\mathbf{u}(t) = \mathbf{f}(t) + \lambda \int \mathbf{H}(t\vartheta) \cdot \mathbf{f}(\vartheta)\, d\vartheta \quad \dots\dots (12).$$

For on multiplying

$$\mathbf{u}(\vartheta) = \mathbf{f}(\vartheta) + \lambda \int \mathbf{K}(\vartheta s) \cdot \mathbf{u}(s)\, ds \quad \dots\dots (1'')$$

by $\lambda \mathbf{H}(t\vartheta)\cdot$ and integrating over the domain we have

$$\lambda \int \mathbf{H}(t\vartheta) \cdot \mathbf{u}(\vartheta)\, d\vartheta = \lambda \int \mathbf{H}(t\vartheta) \cdot \mathbf{f}(\vartheta)\, d\vartheta + \lambda^2 \iint \mathbf{H}(t\vartheta) \cdot \mathbf{K}(\vartheta s) \cdot \mathbf{u}(s)\, ds\, d\vartheta.$$

The factors in the final integral are associative. The order of integration may be changed and the equation then becomes in virtue of (11)

$$0 = \int \mathbf{H}(t\vartheta) \cdot \mathbf{f}(\vartheta)\, d\vartheta - \lambda \int \mathbf{K}(ts) \cdot \mathbf{u}(s)\, ds,$$

which by (1) is equivalent to (12). In particular the series $\mathbf{S}(t)$ of the previous section is the solution of (1) for the restricted values of $\lambda$ there considered. For $\mathbf{S}(t)$ is expressible in the form (9) which is identical with (12), and in which $\mathbf{H}(ts)$ satisfies the resolvent relation (11).

Having shewn that the value of $\mathbf{u}(t)$ given by (12) satisfies the original equation (1), we may prove that this is the only solution. Assume that there is another solution $\mathbf{u}_1(t)$ so that

$$\mathbf{u}_1(t) - \lambda \int \mathbf{K}(ts) \cdot \mathbf{u}_1(s)\, ds = \mathbf{f}(t).$$

Substituting in (12) the value of $\mathbf{f}(t)$ given by this equation we find

$$\mathbf{u}(t) = \mathbf{u}_1(t) - \lambda \int [\mathbf{K}(ts) - \mathbf{H}(ts) + \lambda \int \mathbf{H}(t\vartheta) \cdot \mathbf{K}(\vartheta s)\, d\vartheta] \cdot \mathbf{u}_1(s)\, ds.$$

The integrand vanishes in virtue of (11), shewing that

$$\mathbf{u}(t) = \mathbf{u}_1(t).$$

The solution given by (12) is therefore *unique*. The same equation also puts in evidence the appropriateness of the term "resolvent" as applied to the dyadic $\mathbf{H}(ts)$. As an alternative name we might suggest "solving dyadic."

§ 8. But further, if there exists such a dyadic $\mathbf{H}(ts)$ satisfying (11) it must also satisfy (11′), and conversely. Consider (11′) as an equation whose unknown is $\mathbf{H}(ts)$. This equation is precisely of the form (1). Hence by (12) its unique solution is given by

$$\mathbf{H}(ts) = \mathbf{K}(ts) + \lambda \int \mathbf{H}(t\vartheta) \cdot \mathbf{K}(\vartheta s)\, d\vartheta.$$

This solution is therefore identical with the function $\mathbf{H}(ts)$ given by (11). The relations are therefore reciprocal.

The solution of the *associated equation*

$$\mathbf{v}(t) - \lambda \int \mathbf{v}(s) \cdot \mathbf{K}(st)\, ds = \mathbf{f}(t) \quad \text{.................................(3)}$$

is expressed in terms of the same resolvent $\mathbf{H}(st)$. For on changing $t$ to $\vartheta$ in this equation, multiplying by $\cdot \mathbf{H}(\vartheta t)\lambda d\vartheta$ and integrating, we find exactly as in the previous section that

$$\mathbf{v}(t) = \mathbf{f}(t) + \lambda \int \mathbf{f}(\vartheta) \cdot \mathbf{H}(\vartheta t)\, d\vartheta \quad \text{..............................(13).}$$

That this solution is unique may be shewn as before.

If we put $\mathbf{f}(t) = 0$ in the foregoing, the second members of (12) and (13) disappear. It follows then that *when the kernel* $\mathbf{K}(ts)$ *admits a resolvent* $\mathbf{H}(ts)$ (which will be proved to be the case except for certain isolated values of the parameter $\lambda$) *the homogeneous vector integral equation*

$$\mathbf{u}(t) = \lambda \int \mathbf{K}(t\vartheta) \cdot \mathbf{u}(\vartheta)\, d\vartheta \quad \text{.................................(14)}$$

*and the associated homogeneous equation*

$$\mathbf{v}(t) = \lambda \int \mathbf{v}(\vartheta) \cdot \mathbf{K}(\vartheta t)\, d\vartheta \quad \text{.................................(15)}$$

*have no finite and continuous solutions but* $\mathbf{u}(t) = 0$ *and* $\mathbf{v}(t) = 0$ *respectively.*

In order to remove the restrictions imposed in § 6 on the values of the parameter $\lambda$, and to extend the validity of the solutions (12) and (13) to the case in which the parameter value is quite general, we shall endeavour to find a resolvent dyadic $\mathbf{H}(ts)$ satisfying the relations (11) and (11′), and it will then follow that the unique solutions of the integral equations (1) and (3) will be as expressed in (12) and (13) respectively. If the resolvent as found becomes infinite for any value of $\lambda$, these solutions will in general break down for that value of the parameter. The determination of the resolvent and the examination of its properties will therefore be our main concern.

## IV. Dyadic Determinants.

§ 9. Before however undertaking the search for a resolvent we shall need to introduce and define a class of determinants in which each element is a dyadic. The properties of such will clearly differ in many important respects from those of ordinary determinants where each element is a scalar quantity. The particularly simple properties of scalar determinants are due to the fact that in the expansion the order of the factors is immaterial, a property which is not shared by the factors of a dyadic product. A dyadic determinant of $n^2$ elements may be expanded according to any one of its $n$ rows or of its $n$ columns, and in general each of these methods will yield a different result. And further, in each of these ways of expanding we must adhere to certain definite rules so that the order of the elements in any term will not be arbitrary.

Consider first expansion according to rows. Determinants to be developed in this way we shall speak of as *row-determinants*. That obtained by expanding according to the first row will be the most important for our purpose, and for this first row-determinant we shall adopt the following scheme of expansion. Let $\mathbf{a}_{rs}$, $(r, s = 1, 2, \ldots, n)$ be a set of $n^2$ dyadic elements. The determinant $\begin{vmatrix} \mathbf{a}_{11} & \mathbf{a}_{12} \\ \mathbf{a}_{21} & \mathbf{a}_{22} \end{vmatrix}$ expanded according to the first row is to be interpreted as $(\mathbf{a}_{11} \cdot \mathbf{a}_{22} - \mathbf{a}_{12} \cdot \mathbf{a}_{21})$, the element in the first row being placed first in each term. The determinant

$$\begin{vmatrix} \mathbf{a}_{11} & \mathbf{a}_{12} & \mathbf{a}_{13} \\ \mathbf{a}_{21} & \mathbf{a}_{22} & \mathbf{a}_{23} \\ \mathbf{a}_{31} & \mathbf{a}_{32} & \mathbf{a}_{33} \end{vmatrix}$$

is to mean

$$\mathbf{a}_{11} \cdot \begin{vmatrix} \mathbf{a}_{22} & \mathbf{a}_{23} \\ \mathbf{a}_{32} & \mathbf{a}_{33} \end{vmatrix} - \mathbf{a}_{12} \cdot \begin{vmatrix} \mathbf{a}_{21} & \mathbf{a}_{23} \\ \mathbf{a}_{31} & \mathbf{a}_{33} \end{vmatrix} - \mathbf{a}_{13} \cdot \begin{vmatrix} \mathbf{a}_{31} & \mathbf{a}_{32} \\ \mathbf{a}_{21} & \mathbf{a}_{22} \end{vmatrix}.$$

The leading element $\mathbf{a}_{11}$ is multiplied by its minor in the ordinary sense. The $i$th element in the first row is multiplied by its minor, but in that minor the elements belonging to the $i$th row of the main determinant must occur as the first row. All these products have a negative sign prefixed except that involving the leading element. This method gives a similar result to the expansion of the corresponding scalar determinant, but with our convention as to the order of the elements it will be seen that the second suffix of each element of a product is the same as the first of the next element, until the first suffix in that group of elements is repeated, whereupon the group becomes closed. After that other complete groups may occur. Representative terms from the above expansion are

$$\mathbf{a}_{12} \cdot \mathbf{a}_{23} \cdot \mathbf{a}_{31}, \quad -\mathbf{a}_{13} \cdot \mathbf{a}_{31} \cdot \mathbf{a}_{22}, \quad -\mathbf{a}_{11} \cdot \mathbf{a}_{23} \cdot \mathbf{a}_{32}.$$

In the first of these the closed group embraces all the factors, beginning and ending with the suffix 1. In the second example the first two factors form a group while the third factor forms a complete group by itself. The third example begins with the group $\mathbf{a}_{11}$ which is then followed by another.

Quite generally the dyadic determinant

$$\begin{vmatrix} \mathbf{a}_{11} & \mathbf{a}_{12} & \mathbf{a}_{13} \ldots\ldots \mathbf{a}_{1n} \\ \mathbf{a}_{21} & \mathbf{a}_{22} \ldots\ldots\ldots\ldots & \mathbf{a}_{2n} \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ \mathbf{a}_{n1} & \mathbf{a}_{n2} \ldots\ldots\ldots\ldots & \mathbf{a}_{nn} \end{vmatrix}$$

expanded according to the first row is to be interpreted

$$\mathbf{a}_{11} \bullet \begin{vmatrix} \mathbf{a}_{22} & \mathbf{a}_{23} \ldots\ldots \mathbf{a}_{2n} \\ \mathbf{a}_{32} \ldots\ldots\ldots\ldots & \mathbf{a}_{3n} \\ \ldots\ldots\ldots\ldots\ldots\ldots \\ \mathbf{a}_{n2} \ldots\ldots\ldots\ldots & \mathbf{a}_{nn} \end{vmatrix} - \sum_{i=2}^{n} \mathbf{a}_{1i} \bullet \left[\begin{array}{l}\text{The minor of } \mathbf{a}_{1i} \text{ with the elements} \\ \text{belonging to the } i\text{th row of the} \\ \text{original determinant placed in the} \\ \text{first row}\end{array}\right].$$

In expanding each of these minors the rules laid down for the original determinant are to be observed. The suffixes according to this rule of expansion fall into groups as explained above, each term embracing one or more groups and all the groups being closed.

It is easily verified that the interchange of the first two columns alters the sign of the first row-determinant but leaves it otherwise unchanged. This clearly holds when the determinant consists of $2^2$ or $3^2$ elements, and thence by induction it is shewn to hold generally. It follows then that if the first row-determinant is unaltered by the interchange of the first two columns it must be zero; and therefore, if the first two columns are identical it must be equal to zero. The interchange of rows has no corresponding simple result: for such a change commutes two factors in every term of the expansion, giving in general an entirely different dyadic.

§ 10. Consider next the expansion of *column-determinants.* Expansion according to the first column will be interpreted as follows. The determinant

$$\begin{vmatrix} \mathbf{a}_{11} & \mathbf{a}_{12} \\ \mathbf{a}_{21} & \mathbf{a}_{22} \end{vmatrix} \text{ will mean } (\mathbf{a}_{22} \bullet \mathbf{a}_{11} - \mathbf{a}_{12} \bullet \mathbf{a}_{21}),$$

the element from the first column being placed *last* in each term of the expansion. The determinant

$$\begin{vmatrix} \mathbf{a}_{11} & \mathbf{a}_{12} & \mathbf{a}_{13} \\ \mathbf{a}_{21} & \mathbf{a}_{22} & \mathbf{a}_{23} \\ \mathbf{a}_{31} & \mathbf{a}_{32} & \mathbf{a}_{33} \end{vmatrix}$$

is to be understood as

$$\begin{vmatrix} \mathbf{a}_{22} & \mathbf{a}_{23} \\ \mathbf{a}_{32} & \mathbf{a}_{33} \end{vmatrix} \bullet \mathbf{a}_{11} - \begin{vmatrix} \mathbf{a}_{12} & \mathbf{a}_{13} \\ \mathbf{a}_{32} & \mathbf{a}_{33} \end{vmatrix} \bullet \mathbf{a}_{21} - \begin{vmatrix} \mathbf{a}_{13} & \mathbf{a}_{12} \\ \mathbf{a}_{23} & \mathbf{a}_{22} \end{vmatrix} \bullet \mathbf{a}_{31}$$

in which each of the determinants is to be expanded according to the first column. It will be noticed that the elements from the first column of the original determinant occur last in each product, the prefactor of $\mathbf{a}_{i1}$ being its minor in the main determinant with the elements of the $i$th column of that determinant placed in the first column. Each minor is then to be expanded according to its first column.

The general rule is now clear. The determinant of $n^2$ elements considered above when expanded according to its first column

$$= \begin{vmatrix} \mathbf{a}_{22} & \mathbf{a}_{23} \ldots\ldots & \mathbf{a}_{2n} \\ \mathbf{a}_{32} & \ldots\ldots\ldots & \mathbf{a}_{3n} \\ \ldots & \ldots & \ldots \\ \mathbf{a}_{n2} & \ldots\ldots\ldots & \mathbf{a}_{nn} \end{vmatrix} \bullet \mathbf{a}_{11} - \sum_{i=2}^{n} \begin{bmatrix} \text{The minor of } \mathbf{a}_{i1} \text{ with the elements} \\ \text{belonging to the } i\text{th column of the} \\ \text{original determinant placed in the} \\ \text{first column} \end{bmatrix} \bullet \mathbf{a}_{i1}.$$

Each of these minors is to be expanded according to the first column, the same rules being observed as for the main determinant. In this way the suffixes fall into groups as in the case of row-determinants, each term of the expansion embracing one or more closed groups.

It is easy to shew that the interchange of the first two rows alters the sign of the first column-determinant, leaving it otherwise unchanged. Hence, if the first column-determinant is unaltered by the interchange of the first two rows it must be equal to zero; and thus if the first two rows are identical the first column-determinant vanishes. The interchange of columns gives quite a different dyadic.

## V. The Series $D(\lambda)$.

§ 11. Knowing beforehand that the solution of the single scalar integral equation (5) is a meromorphic function of the parameter $\lambda$, and that the vector integral equation (1) is equivalent to the system (2) which in turn can be replaced by the single integral equation (5), we are prepared to find the solution $\mathbf{u}(t)$ of (1) also a meromorphic function of $\lambda$. We expect then to determine the resolvent $\mathbf{H}(ts)$ of this integral equation as the quotient of two integral functions of $\lambda$, the roots of the denominator being the singular parameter values for which the solution in general breaks down.

In the search for these integral functions we might adopt the method of replacing the integral equation (1) by a set of $n$ vector algebraical equations with $n$ unknowns*, and proceed to the limit $n = \infty$. But it is perhaps easier to pursue a different path. The series known as the "determinant" of the scalar equation (5) should be identical with that which is to furnish the characteristic numbers of the vector equation (1). We shall therefore consider this infinite series, and if possible express it in terms of the dyadic $\mathbf{K}(ts)$. But in order not to assume that this series in the new form does furnish the singular values of the parameter $\lambda$, we shall merely take it as one suggested by previous knowledge, and seek to determine a meromorphic function of $\lambda$ whose denominator is this suggested series, and which satisfies the resolvent relations (11) and (11′). If such a function can be found it will follow that the roots of its denominator are the singular parameter values for the solution $\mathbf{u}(t)$.

Now the infinite series whose roots are the characteristic numbers of the equation (5) is known to be

$$1 - \lambda \int_{S'} K\begin{pmatrix} s_1 \\ s_1 \end{pmatrix} ds_1 + \frac{\lambda^2}{2\,!} \int_{S'} K\begin{pmatrix} s_1 s_2 \\ s_1 s_2 \end{pmatrix} ds_1 ds_2 - \frac{\lambda^3}{3\,!} \int_{S'} K\begin{pmatrix} s_1 s_2 s_3 \\ s_1 s_2 s_3 \end{pmatrix} ds_1 ds_2 ds_3 + \ldots \quad \ldots\ldots(16),$$

* This set of $n$ vector equations could then be replaced by $3n$ scalar equations. In order that this system may admit a solution the determinant of the $(3n)^2$ coefficients must not vanish. If $n$ becomes infinitely great and this determinant is expanded in powers of $\lambda$, we recognise, though perhaps with some difficulty, the same series (18) below.

where, according to the usual notation*,

$$K\begin{pmatrix} s_1 s_2 \dots s_n \\ t_1 t_2 \dots t_n \end{pmatrix} \equiv \begin{vmatrix} K(s_1t_1) & K(s_1t_2) & \dots & K(s_1t_n) \\ K(s_2t_1) & K(s_2t_2) & \dots & K(s_2t_n) \\ \dots & \dots & \dots & \dots \\ K(s_nt_1) & K(s_nt_2) & \dots & K(s_nt_n) \end{vmatrix} \qquad \dots\dots(17)$$

is an ordinary scalar determinant. The suffix $S'$ to the integral sign in (16) indicates that the region of integration is $S'$, defined in § 4, while $K(st)$ is equal successively to the functions $K_{nm}(st)$ $(n, m = 1, 2, 3)$ as explained in that section. Expanding the determinants, and replacing the multiple domain of integration by the original domain $S$, we find for the earlier terms of the series

$$1 - \lambda \int [K_{11}(s_1s_1) + K_{22}(s_1s_1) + K_{33}(s_1s_1)]\, ds_1$$
$$+ \frac{\lambda^2}{2!} \iint [\{K_{11}(s_1s_1) + K_{22}(s_1s_1) + K_{33}(s_1s_1)\}\{K_{11}(s_2s_2) + K_{22}(s_2s_2) + K_{33}(s_2s_2)\}$$
$$- \{K_{11}(s_1s_2) K_{11}(s_2s_1) + K_{12}(s_1s_2) K_{21}(s_2s_1) + K_{13}(s_1s_2) K_{31}(s_2s_1)$$
$$+ K_{21}(s_1s_2) K_{12}(s_2s_1) + \dots\dots$$
$$+ K_{31}(s_1s_2) K_{13}(s_2s_1) + \dots\dots\dots\dots + K_{33}(s_1s_2) K_{33}(s_2s_1)\}]\, ds_1 ds_2.$$

Now $K_{11}(ts) + K_{22}(ts) + K_{33}(ts)$ is the first scalar or briefly the scalar of the dyadic $\mathbf{K}(ts)$ and is usually denoted by $\mathbf{K}_S(ts)$. It is an invariant† of the dyadic equal to the sum of the coefficients in the main diagonal when the dyadic is expressed in nonion form. Thus the coefficient of $\lambda$ is equal to $\int \mathbf{K}_S(s_1s_1)\, ds_1$. Similarly in the coefficient of $\lambda^2$ we have under the integral sign first the product of the scalars of $\mathbf{K}(s_1s_1)$ and $\mathbf{K}(s_2s_2)$, followed by an expression which is easily recognised as the scalar of the product‡ $\mathbf{K}(s_1s_2) \cdot \mathbf{K}(s_2s_1)$. The coefficient of $\lambda^2/2!$ could therefore be written

$$\iint \begin{vmatrix} \mathbf{K}_S(s_1s_1) & \mathbf{K}(s_1s_2) \\ \mathbf{K}(s_2s_1) & \mathbf{K}_S(s_2s_2) \end{vmatrix} ds_1 ds_2,$$

provided we interpret $\iint \mathbf{K}(s_1s_2) \cdot \mathbf{K}(s_2s_1)\, ds_2 ds_1$ as meaning

$$\int ds_1 [\int \mathbf{K}(s_1s_2) \cdot \mathbf{K}(s_2s_1)\, ds_2]_S \equiv \int \mathbf{K}_{1S}(tt)\, dt,$$

that is, the integral of the scalar of $\mathbf{K}_1(tt)$. Similarly on expanding the coefficient of $\lambda^3/3!$ we should find an expression which is identical with the determinant

$$\iiint \begin{vmatrix} \mathbf{K}_S(s_1s_1) & \mathbf{K}(s_1s_2) & \mathbf{K}(s_1s_3) \\ \mathbf{K}(s_2s_1) & \mathbf{K}_S(s_2s_2) & \mathbf{K}(s_2s_3) \\ \mathbf{K}(s_3s_1) & \mathbf{K}(s_3s_2) & \mathbf{K}_S(s_3s_3) \end{vmatrix} ds_1 ds_2 ds_3$$

expanded according to either the first row or the first column, *provided we interpret the integrals as follows*:

$$\int \mathbf{K}(s_1s_1)\, ds_1 \equiv \int \mathbf{K}_S(s_1s_1)\, ds_1,$$
$$\iint \mathbf{K}(s_1s_2) \cdot \mathbf{K}(s_2s_1)\, ds_2 ds_1 \equiv \int \mathbf{K}_{1S}(s_1s_1)\, ds_1,$$
$$\iiint \mathbf{K}(s_1s_2) \cdot \mathbf{K}(s_2s_3) \cdot \mathbf{K}(s_3s_1)\, ds_1 ds_2 ds_3 \equiv \int \mathbf{K}_{2S}(s_1s_1)\, ds_1,$$
$$\iiint \mathbf{K}(s_1s_2) \cdot \mathbf{K}(s_2s_1) \cdot \mathbf{K}(s_3s_3)\, ds_1 ds_2 ds_3 \equiv \int \mathbf{K}_{1S}(s_1s_1)\, \mathbf{K}_S(s_3s_3)\, ds_1 ds_3.$$

* Fredholm, *loc. cit.* p. 367.

† Gibbs-Wilson, *loc. cit.* p. 319.

‡ *Ibid.* p. 318.

and so on, where $\mathbf{K}_{i_S}(tt)$ denotes the scalar of the iterated dyadic $\mathbf{K}_i(tt)$. With this convention it is immaterial whether we retain the suffix $S$ in the leading diagonal of the above determinants. According to the above scheme of interpretation, as soon as a group of factors in any term becomes closed, that is as soon as the initial variable of integration is repeated, the integral of the dyadic product in that closed group is to be replaced by the integral of the scalar of the corresponding iterated dyadic. If any term involves more than one group the integral of that term is to be understood as the integral of the product of the scalars of those closed groups; or in other words (since the variables in the separate scalars are different), as the product of the integrals of the scalars of the closed groups. The above determinants are therefore scalar quantities. Each term of the expansion is either the integral of the scalar of an iterated dyadic, or the product of two or more such integrals.

§ 12. We might stop to prove that the coefficients of the successive powers of $\lambda$ fall into determinants of scalars as above, and that therefore the new scalar series is absolutely convergent because it is identical with the absolutely convergent series (16). But this course is unnecessary. We have simply used the known series (16) to *suggest* the corresponding series for the equation (1). We shall take the series suggested and prove that it satisfies the requirements of the problem.

Consider then the series suggested, viz.

$$D(\lambda) = 1 - \lambda \int \mathbf{K}_S(s_1 s_1)\, ds_1 + \frac{\lambda^2}{2!} \int\int \begin{vmatrix} \mathbf{K}_S(s_1 s_1) & \mathbf{K}(s_1 s_2) \\ \mathbf{K}(s_2 s_1) & \mathbf{K}_S(s_2 s_2) \end{vmatrix} ds_1 ds_2$$

$$- \ldots + (-)^n \frac{\lambda^n}{n!} \int \ldots \int \mathbf{K}\begin{pmatrix} s_1 s_2 \ldots s_n \\ s_1 s_2 \ldots s_n \end{pmatrix} ds_1 \ldots ds_n + \ldots \qquad \text{(18)},$$

where in the general term we have adopted the notation (17). The integrals obtained by expanding these determinants are to be interpreted as explained in the previous section. It is clear upon examination that the determinants may be expanded as either first row or first column determinants, the same result being obtained by either method. For the factors in any term of the expansion fall into closed groups in just the same way, the only difference being that the order of the closed groups in any term will be different. But as the integral of the closed group represents a scalar quantity, the change of order is immaterial.

The series $D(\lambda)$ is absolutely convergent. An upper limit may be assigned to the value of the integral which is the coefficient of $(-\lambda)^n/n!$. In accordance with the meaning of the expression "greatest modular magnification," $M$, of § 5, it is clear that $\mathbf{K}(ts) \leqslant M\mathbf{I}$, meaning that the modulus of the vector after operation with $\mathbf{K}(ts)$ is $\leqslant$ that after operation with $M\mathbf{I}$. Each element of the determinant regarded as a dyadic cannot exceed $M\mathbf{I}$. Further, the scalar of any dyadic which $\leqslant p\mathbf{I}$ cannot exceed $pN$, where $N$ is a finite scalar quantity which may be taken greater than unity. In any term of the expansion of the determinant there are not more than $n$ groups; hence the determinant is equivalent to a scalar determinant the absolute value of each term of which does not exceed

$MN$. If $B$ is the measure of the region of integration the general term of (18) therefore does not exceed

$$C_n \equiv \sqrt{n^n}\, \frac{(MNB)^n \lambda^n}{n!}$$

in virtue of Hadamard's theorem*. The ratio of consecutive terms of the series whose general term is $C_n$ is

$$\frac{C_{n+1}}{C_n} = \frac{MNB\lambda}{\sqrt{n+1}} \left(1 + \frac{1}{n}\right)^{\frac{n}{2}}$$

which tends to zero as $n$ tends to infinity. The series $\sum_n^{\infty} C_n$ is thus absolutely convergent. Therefore the series $D(\lambda)$, whose general term is numerically $\leqslant C_n$, must also be absolutely convergent.

## VI. Determination of Adjoint and Resolvent.

§ 13. Let us now try to find, if possible, a meromorphic function $\mathbf{H}(ts:\lambda)$ of $\lambda$, satisfying (11), having this series $D(\lambda)$ as denominator and another integral function $\mathbf{D}(ts:\lambda)$ as numerator, so that

$$\mathbf{H}(ts:\lambda) = \frac{\mathbf{D}(ts:\lambda)}{D(\lambda)} \quad \text{.......................................(19).}$$

On substitution of this value the equation (11) becomes

$$\mathbf{D}(ts:\lambda) - D(\lambda)\,\mathbf{K}(ts) = \lambda \int \mathbf{D}(t\vartheta:\lambda) \cdot \mathbf{K}(\vartheta s)\, d\vartheta \quad \text{.....................(20).}$$

If the numerator in the second member of (19) is an integral function of $\lambda$ it may be written

$$\mathbf{D}(ts:\lambda) = \mathbf{A}_0(ts) - \lambda \mathbf{A}_1(ts) + \lambda^2 \mathbf{A}_2(ts) - \ldots \quad \text{.....................(21),}$$

where the coefficients $\mathbf{A}_n(ts)$ are a series of dyadic functions of the positions of the points $t$ and $s$. We wish, if possible, to determine these coefficients so that the relation (20) may hold. Let $a_n$ denote the coefficient of $(-\lambda)^n$ in (18). Then on substituting in (20) the values of the functions given by (18) and (21), and equating coefficients of the different powers of $\lambda$, we have the following relations

$$\left.\begin{aligned} &\mathbf{A}_0(ts) = \mathbf{K}(ts) \\ &\mathbf{A}_1(ts) = a_1 \mathbf{K}(ts) - \int \mathbf{A}_0(t\vartheta) \cdot \mathbf{K}(\vartheta s)\, d\vartheta \\ &\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ &\mathbf{A}_n(ts) = a_n \mathbf{K}(ts) - \int \mathbf{A}_{n-1}(t\vartheta) \cdot \mathbf{K}(\vartheta s)\, d\vartheta \\ &\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \end{aligned}\right\} \quad \text{.....................(22).}$$

Substituting the values of the coefficients $a_n$ we find

$$\mathbf{A}_1(ts) = \int \begin{vmatrix} \mathbf{K}(ts) & \mathbf{K}(ts_1) \\ \mathbf{K}(s_1 s) & \mathbf{K}_S(s_1 s_1) \end{vmatrix} ds_1,$$

$$2\mathbf{A}_2(ts) = 2a_2 \mathbf{K}(ts) - \int \mathbf{A}_1(ts_1) \cdot \mathbf{K}(s_1 s)\, ds_1 - \int \mathbf{A}_1(ts_2) \cdot \mathbf{K}(s_2 s)\, ds_2$$

$$= \iint \begin{vmatrix} \mathbf{K}(ts) & \mathbf{K}(ts_1) & \mathbf{K}(ts_2) \\ \mathbf{K}(s_1 s) & \mathbf{K}_S(s_1 s_1) & \mathbf{K}(s_1 s_2) \\ \mathbf{K}(s_2 s) & \mathbf{K}(s_2 s_1) & \mathbf{K}_S(s_2 s_2) \end{vmatrix} ds_1 ds_2.$$

* *Bulletin des Sciences Math.* (2), Vol. XVII. (1893), pp. 240–2.

In interpreting these determinants the same convention is to be observed as in § 11. Whenever in any dyadic product the group becomes closed by the recurrence of the initial variable in that group, the integral is to be replaced by the integral of the scalar of the appropriate iterated dyadic. But each of the above determinants is a dyadic. There are in every term two or more consecutive factors not forming a closed group, e.g.

$$\int \mathbf{K}(ts_2) \cdot \mathbf{K}(s_2 s)\, ds_2,$$

which is the iterated dyadic $\mathbf{K}_1(ts)$. Hence every term of the expansion is a dyadic. The determinants may be expanded according to either the first row or the first column. The result is the same in each case, for the unclosed dyadic factor occurs in the one case before, and in the other after the scalar factors in that term.

Proceeding in the same way we find that

$$\mathbf{A}_3(ts) = \frac{1}{3!}\iiint \mathbf{K}\begin{pmatrix} ts_1s_2s_3 \\ ss_1s_2s_3 \end{pmatrix} ds_1 ds_2 ds_3.$$

The law of formation of these coefficients is quite general. Assume that it holds for $\mathbf{A}_{n-1}(ts)$. Then since

$$na_n = \int \mathbf{A}_{n-1}(s_1s_1)\, ds_1$$

we have by (22)

$$n\mathbf{A}_n(ts) = \mathbf{K}(ts)\int \mathbf{A}_{n-1}(s_1s_1)\, ds_1 - n\int \mathbf{A}_{n-1}(t\vartheta) \cdot \mathbf{K}(\vartheta s)\, d\vartheta.$$

In the last term of this equation we may drop the factor $n$ and take the sum of $n$ separate integrals in which $\vartheta$ is replaced successively by $s_1, s_2, \ldots s_n$. On substitution of the value of $\mathbf{A}_{n-1}(ts)$ the equation then gives immediately

$$\mathbf{A}_n(ts) = \frac{1}{n!}\int \ldots \int \begin{vmatrix} \mathbf{K}(ts) & \mathbf{K}(ts_1) & \ldots\ldots & \mathbf{K}(ts_n) \\ \mathbf{K}(s_1s) & \mathbf{K}(s_1s_1) & \ldots\ldots & \mathbf{K}(s_1s_n) \\ \ldots & \ldots & \ldots & \ldots \\ \mathbf{K}(s_ns) & \mathbf{K}(s_ns_1) & \ldots\ldots & \mathbf{K}(s_ns_n) \end{vmatrix} ds_1 \ldots ds_n.$$

We have thus found a solution of the equation (11) in the form of a meromorphic function $\mathbf{H}(ts:\lambda)$ of $\lambda$ given by (19). The denominator $D(\lambda)$ is the infinite series (18), and the numerator the infinite series just found, viz.

$$\mathbf{D}(ts:\lambda) = \mathbf{K}(ts) - \lambda\int \mathbf{K}\begin{pmatrix} ts_1 \\ ss_1 \end{pmatrix} ds_1 + \ldots + \frac{(-\lambda)^n}{n!}\int \ldots \int \mathbf{K}\begin{pmatrix} ts_1s_2 \ldots s_n \\ ss_1s_2 \ldots s_n \end{pmatrix} ds_1 \ldots ds_n + \ldots \quad \ldots (23),$$

which may be called the *adjoint* of $\mathbf{K}(ts)$.

This series is absolutely and uniformly convergent in the region considered. In the general term the coefficient of $(-\lambda)^{n-1}/(n-1)!$ is a determinant which represents a dyadic. The elements of the first row and the first column are each, in the notation of the previous section, $\leqslant M\mathbf{I}$. In any term of the expanded determinant there are not more than $(n-1)$ closed groups. The number of variables of integration is $(n-1)$: so that, in virtue of Hadamard's theorem already mentioned, the absolute value of the integrated determinant

$$\leqslant \sqrt{n^n}\, M^n (BX)^{n-1}\, \mathbf{I}.$$

The general term of the series (23) therefore does not exceed in value the dyadic

$$\mathbf{C}_n = \frac{\lambda^{n-1}}{(n-1)!}\sqrt{n^n}\, M^n (BX)^{n-1}\, \mathbf{I}.$$

The ratio of consecutive terms of the series $\sum_{n}^{\infty} \mathbf{C}_n$ is

$$\frac{\mathbf{C}_{n+1}}{\mathbf{C}_n} = \frac{\lambda MBN}{\sqrt{n}} \left(1 + \frac{1}{n}\right)^{\frac{n+1}{2}},$$

which tends to zero as $n$ tends to infinity. The series $\sum_{n}^{\infty} \mathbf{C}_n$ is thus absolutely convergent, and therefore the series (23) absolutely and uniformly convergent. That the series is also continuous within $S$ at every point at which $\mathbf{K}(ts)$ is continuous may be established by the same argument as for the scalar integral equation*.

§ 14. The quotient of the two integral functions $\mathbf{D}(ts:\lambda)$ and $D(\lambda)$ is the resolvent $\mathbf{H}(ts:\lambda)$ for which we have been enquiring; for in the preceding section the first of these functions was determined so that the quotient would satisfy (11) and therefore (11'). The solution of the integral equation (1) is then given by (12) for all values of $\lambda$ for which $\mathbf{H}(ts:\lambda)$ has a meaning. Now the only parameter values for which it can cease to have a meaning are the roots of the denominator $D(\lambda)$. These roots we shall term the "singular parameter values" or the "characteristic numbers." That every characteristic number is actually a pole of the resolvent may be proved thus. Suppose $\lambda_0$ is a root of $D(\lambda)$ of multiplicity $p$. Then

$$\frac{d^i}{d\lambda_0{}^i} D(\lambda_0) = 0, \quad i = 1, 2, \ldots, (p-1)$$

while

$$\frac{d^p}{d\lambda_0{}^p} D(\lambda_0) \neq 0.$$

But on comparison of the series for $D(\lambda)$ and $\mathbf{D}(ts:\lambda)$ it is clear that

$$\frac{d^p}{d\lambda^p} D(\lambda) = \frac{d^{p-1}}{d\lambda^{p-1}} \int \mathbf{D}(ss:\lambda)\, ds \quad \ldots\ldots\ldots\ldots(24),$$

where the integral in the second member is as usual to be interpreted as the integral of the scalar $\mathbf{D}_S(ss:\lambda)$. Now as the first member of (24) does not vanish when $\lambda = \lambda_0$, the second cannot. But when the scalar of a dyadic does not vanish the dyadic itself cannot vanish identically: so that

$$\frac{d^{p-1}}{d\lambda_0{}^{p-1}} \mathbf{D}(ts:\lambda_0) \neq 0.$$

It follows then that the multiplicity of $\lambda_0$ regarded as a root of the adjoint is at least one less than $p$. The value $\lambda_0$ is therefore a pole of the resolvent. The solution (12) of our integral equation (1) thus breaks down at a singular parameter value, unless conditions are satisfied which neutralize the effect of the pole of the resolvent.

## VII. THE HOMOGENEOUS INTEGRAL EQUATIONS, AND SINGULAR PARAMETER VALUES.

§ 15. It was shewn in § 8 that when the resolvent exists neither the homogeneous equation

$$\mathbf{u}(t) = \lambda \int \mathbf{K}(ts) \cdot \mathbf{u}(s)\, ds \quad \ldots\ldots\ldots\ldots(14)$$

nor its associated equation

$$\mathbf{v}(t) = \lambda \int \mathbf{v}(s) \cdot \mathbf{K}(st)\, ds \quad \ldots\ldots\ldots\ldots(15)$$

* Cf. e.g. Bôcher, "An Introduction to the study of Integral Equations," *Cambridge Tract* (1914), pp. 33–35.

admits any finite and continuous solution but zero. We shall now prove that *when* $\lambda$ *is equal to a characteristic number* $\lambda_0$ *each of these equations admits at least one non-zero solution.*

Since $\lambda_0$ is a pole of the resolvent, this dyadic may in the neighbourhood of $\lambda_0$ be expressed in the form

$$\mathbf{H}(ts:\lambda) = \frac{\mathbf{B}_r(ts)}{(\lambda_0-\lambda)^r} + \frac{\mathbf{B}_{r-1}(ts)}{(\lambda_0-\lambda)^{r-1}} + \ldots + \mathbf{B}_0(ts:\lambda) \quad \ldots\ldots\ldots\ldots\ldots(25),$$

$r$ being the order of the pole, $\mathbf{B}_i(ts)$, $(i = 1, 2, \ldots r)$, a set of dyadic functions of the positions of $t$ and $s$, and $\mathbf{B}_0(ts:\lambda)$ a dyadic holomorphic in $\lambda$. If we substitute this value of the resolvent in the equation (11) written in the form

$$\mathbf{H}(ts) - \mathbf{K}(ts) = (\lambda-\lambda_0)\int \mathbf{H}(t\vartheta)\bullet\mathbf{K}(\vartheta s)\,d\vartheta + \lambda_0\int\mathbf{H}(t\vartheta)\bullet\mathbf{K}(\vartheta s)\,d\vartheta$$

and equate coefficients of $(\lambda_0-\lambda)^{-r}$ and $(\lambda_0-\lambda)^{-r+1}$ we find

$$\left.\begin{array}{ll}(a) & \mathbf{B}_r(ts) = \lambda_0\int\mathbf{B}_r(t\vartheta)\bullet\mathbf{K}(\vartheta s)\,d\vartheta \\ (b) & \mathbf{B}_{r-1}(ts) = \lambda_0\int\mathbf{B}_{r-1}(t\vartheta)\bullet\mathbf{K}(\vartheta s)\,d\vartheta - \int\mathbf{B}_r(t\vartheta)\bullet\mathbf{K}(\vartheta s)\,d\vartheta\end{array}\right\} \ldots\ldots\ldots\ldots(26).$$

Similarly on substituting from (25) in the alternative relation (11′) and equating coefficients we have, among others, the equations

$$\left.\begin{array}{ll}(a) & \mathbf{B}_r(ts) = \lambda_0\int\mathbf{K}(t\vartheta)\bullet\mathbf{B}_r(\vartheta s)\,d\vartheta \\ (b) & \mathbf{B}_{r-1}(ts) = \lambda_0\int\mathbf{K}(t\vartheta)\bullet\mathbf{B}_{r-1}(\vartheta s)\,d\vartheta - \int\mathbf{K}(t\vartheta)\bullet\mathbf{B}_r(\vartheta s)\,d\vartheta\end{array}\right\} \ldots\ldots\ldots\ldots(27).$$

The relations (26 $a$) and (27 $a$) shew that the dyadic $\mathbf{B}_r(ts)$ regarded as a function of $t$ is a solution of the equation (14), and that $\mathbf{B}_r(st)$ regarded as a function of $t$ is a solution of the associated equation (15). These dyadic solutions may be replaced by ordinary vector solutions: for if $\mathbf{a}$ is any vector quantity it follows from the preceding that $\mathbf{B}_r(ts)\bullet\mathbf{a}$ is a solution in $t$ of (14): and that $\mathbf{a}\bullet\mathbf{B}_r(st)$ is a solution in $t$ of the associated equation.

§ 16. We shall now find expressions for the solutions of the homogeneous equations (14) and (15) at a singular value of $\lambda$ in terms of the row and column dyadic determinants previously introduced. A small prefix $r$ will be used to denote that the determinants involved are to be expanded according to the first row: while the prefix $c$ will indicate column expansion. Thus the expressions

$${}_r\mathbf{K}\begin{pmatrix}s_1\, s_2 \ldots s_n \\ t_1\, t_2 \ldots t_n\end{pmatrix} \text{ and } {}_c\mathbf{K}\begin{pmatrix}s_1\, s_2 \ldots s_n \\ t_1\, t_2 \ldots t_n\end{pmatrix}$$

denote the dyadic determinant of the form (17) expanded according to the first row and the first column respectively.

Introducing series analogous to Fredholm's minors we shall speak of

$${}_r\mathbf{D}_n\begin{pmatrix}s_1\, s_2 \ldots s_n \\ t_1\, t_2 \ldots t_n\end{pmatrix}\lambda\Big) \equiv {}_r\mathbf{K}\begin{pmatrix}s_1\, s_2 \ldots s_n \\ t_1\, t_2 \ldots t_n\end{pmatrix} - \lambda\int {}_r\mathbf{K}\begin{pmatrix}s_1 \ldots s_n\, \tau_1 \\ t_1 \ldots t_n\, \tau_1\end{pmatrix} d\tau_1$$
$$+ \ldots + \frac{(-\lambda)^m}{m!}\int\ldots\int {}_r\mathbf{K}\begin{pmatrix}s_1 \ldots s_n\, \tau_1 \ldots \tau_m \\ t_1 \ldots t_n\, \tau_1 \ldots \tau_m\end{pmatrix} d\tau_1 \ldots d\tau_m + \ldots \quad \ldots\ldots\ldots(28)$$

as the $n$th row-minor of the series $D(\lambda)$: and of

$${}_c\mathbf{D}_n\begin{pmatrix}s_1\, s_2 \ldots s_n \\ t_1\, t_2 \ldots t_n\end{pmatrix}\lambda\Big) \equiv \text{a similar expression in column determinants} \ldots\ldots\ldots\ldots(29)$$

as the $n$th column-minor of $D(\lambda)$. From these definitions it is clear that

$$\frac{d^n}{d\lambda^n} D(\lambda) = (-)^n \int \ldots \int {}_r\mathbf{D}_n \begin{pmatrix} s_1 \ldots s_n \\ s_1 \ldots s_n \end{pmatrix} \lambda \Big) ds_1 \ldots ds_n$$

$$= (-)^n \int \ldots \int {}_c\mathbf{D}_n \begin{pmatrix} s_1 \ldots s_n \\ s_1 \ldots s_n \end{pmatrix} \lambda \Big) ds_1 \ldots ds_n \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots (30),$$

the integrated determinants giving the same result whether expanded according to the first row or the first column.

Developing the terms of (28) according to the first row we have

$$\int \ldots \int {}_r\mathbf{K} \begin{pmatrix} s_1 \ldots s_n \, \tau_1 \ldots \tau_m \\ t_1 \ldots t_n \, \tau_1 \ldots \tau_m \end{pmatrix} d\tau_1 \ldots d\tau_m = \mathbf{K}(s_1 t_1) \cdot \int \ldots \int {}_r\mathbf{K} \begin{pmatrix} s_2 \ldots s_n \, \tau_1 \ldots \tau_m \\ t_2 \ldots t_n \, \tau_1 \ldots \tau_m \end{pmatrix} d\tau_1 \ldots d\tau_m$$

$$- \sum_{i=2}^{n} \mathbf{K}(s_1 t_i) \cdot \int \ldots \int {}_r\mathbf{K} \begin{pmatrix} s_i s_2 \ldots s_{i-1} s_{i+1} \ldots s_n \, \tau_1 \ldots \tau_m \\ t_1 t_2 \ldots t_{i-1} t_{i+1} \ldots t_n \, \tau_1 \ldots \tau_m \end{pmatrix} d\tau_1 \ldots d\tau_m$$

$$- \sum_{i=1}^{m} \int \mathbf{K}(s_1 \tau_i) \cdot \int \ldots \int {}_r\mathbf{K} \begin{pmatrix} \tau_i s_2 \ldots s_n \, \tau_1 \ldots \tau_{i-1} \tau_{i+1} \ldots \tau_m \\ t_1 t_2 \ldots t_n \, \tau_1 \ldots \tau_{i-1} \tau_{i+1} \ldots \tau_m \end{pmatrix} d\tau_1 \ldots d\tau_m \quad \ldots\ldots (31).$$

The last summation may clearly be written

$$- m \int \mathbf{K}(s_1 \tau) \cdot \int \ldots \int {}_r\mathbf{K} \begin{pmatrix} \tau \, s_2 \ldots s_n \, \tau_1 \ldots \tau_{m-1} \\ t_1 \, t_2 \ldots t_n \, \tau_1 \ldots \tau_{m-1} \end{pmatrix} d\tau_1 \ldots d\tau_{m-1} \, d\tau.$$

Multiplying then both sides of (31), thus transformed, by $(-\lambda)^m/m!$ and summing for all values of $m$ from 0 to $\infty$ as required by (28) we derive

$${}_r\mathbf{D}_n \begin{pmatrix} s_1 s_2 \ldots s_n \\ t_1 t_2 \ldots t_n \end{pmatrix} \lambda \Big) = \mathbf{K}(s_1 t_1) \cdot {}_r\mathbf{D}_{n-1} \begin{pmatrix} s_2 \ldots s_n \\ t_2 \ldots t_n \end{pmatrix} \lambda \Big)$$

$$- \sum_{i=2}^{n} \mathbf{K}(s_1 t_i) \cdot {}_r\mathbf{D}_{n-1} \begin{pmatrix} s_i s_2 \ldots s_{i-1} s_{i+1} \ldots s_n \\ t_1 t_2 \ldots t_{i-1} t_{i+1} \ldots t_n \end{pmatrix} \lambda \Big)$$

$$+ \lambda \int \mathbf{K}(s_1 \tau) \cdot {}_r\mathbf{D}_n \begin{pmatrix} \tau \, s_2 \ldots s_n \\ t_1 t_2 \ldots t_n \end{pmatrix} \lambda \Big) d\tau \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (32).$$

Similarly by expanding the column determinants in (29) we find

$${}_c\mathbf{D}_n \begin{pmatrix} s_1 s_2 \ldots s_n \\ t_1 t_2 \ldots t_n \end{pmatrix} \lambda \Big) = {}_c\mathbf{D}_{n-1} \begin{pmatrix} s_2 \ldots s_n \\ t_2 \ldots t_n \end{pmatrix} \lambda \Big) \cdot \mathbf{K}(s_1 t_1)$$

$$- \sum_{i=2}^{n} {}_c\mathbf{D}_{n-1} \begin{pmatrix} s_1 s_2 \ldots s_{i-1} s_{i+1} \ldots s_n \\ t_i t_2 \ldots t_{i-1} t_{i+1} \ldots t_n \end{pmatrix} \lambda \Big) \cdot \mathbf{K}(s_i t_1)$$

$$+ \lambda \int {}_c\mathbf{D}_n \begin{pmatrix} s_1 s_2 \ldots s_n \\ \tau t_2 \ldots t_n \end{pmatrix} \lambda \Big) \cdot \mathbf{K}(\tau t_1) \, d\tau \quad \ldots\ldots\ldots\ldots\ldots\ldots (33).$$

Now since $D(\lambda)$ is an integral function of $\lambda$ not vanishing identically it follows that

$$D(\lambda), \; \frac{d}{d\lambda} D(\lambda), \; \ldots, \; \frac{d^i}{d\lambda^i} D(\lambda), \; \ldots$$

are not all zero when $\lambda = \lambda_0$; and therefore in virtue of (30) that the dyadics

$$\mathbf{D}_1 \begin{pmatrix} s_1 \\ t_1 \end{pmatrix} \lambda_0 \Big), \quad \mathbf{D}_2 \begin{pmatrix} s_1 s_2 \\ t_1 t_2 \end{pmatrix} \lambda_0 \Big), \quad \mathbf{D}_3 \begin{pmatrix} s_1 s_2 s_3 \\ t_1 t_2 t_3 \end{pmatrix} \lambda_0 \Big), \quad \ldots \quad \ldots\ldots\ldots\ldots\ldots\ldots (34)$$

are not all identically zero with respect to the variables. Let $q$ be the index of the first of these which is not identically zero. This number $q$ will be the same whether the determinants are row or column determinants. For these differ only by an interchange of variables, and the identical vanishing of ${}_r\mathbf{D}_n(\ldots)$ involves that of ${}_c\mathbf{D}_n(\ldots)$ and vice versa. It follows then from (32) and (33), since all the minors of index $< q$ vanish identically, that

$$ {}_r\mathbf{D}_q\begin{pmatrix} s_1 s_2 \ldots s_q \\ t_1 t_2 \ldots t_q \end{pmatrix}\lambda_0\Big) = \lambda_0 \int \mathbf{K}(s_1\tau) \bullet {}_r\mathbf{D}_q\begin{pmatrix} \tau s_2 \ldots s_q \\ t_1 t_2 \ldots t_q \end{pmatrix}\lambda_0\Big)\, d\tau \quad \ldots\ldots(35),$$

and

$$ {}_c\mathbf{D}_q\begin{pmatrix} s_1 s_2 \ldots s_q \\ t_1 t_2 \ldots t_q \end{pmatrix}\lambda_0\Big) = \lambda_0 \int {}_c\mathbf{D}_q\begin{pmatrix} s_1 s_2 \ldots s_q \\ \tau t_2 \ldots t_q \end{pmatrix}\lambda_0\Big) \bullet \mathbf{K}(\tau t_1)\, d\tau \quad \ldots\ldots(36).$$

Since we may choose the quantities $s_i, t_i$ $(i = 1, 2, \ldots q)$ so that the minors do not vanish identically, the relations just established shew that the row-minor

$$ {}_r\mathbf{D}_q\begin{pmatrix} s\, s_2 \ldots s_q \\ t_1 t_2 \ldots t_q \end{pmatrix}\lambda_0\Big),$$

regarded as a function of $s$, is a non-zero solution of the homogeneous integral equation (14) for the characteristic number $\lambda_0$; and that the column-minor

$$ {}_c\mathbf{D}_q\begin{pmatrix} s_1 s_2 \ldots s_q \\ t\ \ t_2 \ldots t_q \end{pmatrix}\lambda_0\Big),$$

regarded as a function of $t$, is a non-zero solution of the associated homogeneous equation (15). These solutions are of course dyadics. A vector solution may be obtained in the first case by operating on any vector with the row-minor as a prefactor; and in the second case by operating with the column-minor as a postfactor.

By adopting rules corresponding to those of § 9 for expanding the above determinants according to the $i$th row or the $i$th column, it may be shewn that there are $q$ linearly independent solutions

$$\mathbf{u}_1(t),\ \mathbf{u}_2(t),\ \ldots,\ \mathbf{u}_q(t) \quad \ldots\ldots(37),$$

to the homogeneous integral equation (14) for the singular parameter value $\lambda_0$, and $q$ also

$$\mathbf{v}_1(t),\ \mathbf{v}_2(t),\ \ldots,\ \mathbf{v}_q(t) \quad \ldots\ldots(38),$$

to the associated equation (15) for the same value of $\lambda$. The proof however will not be given. Sufficient has been done to shew that with the aid of our dyadic determinants the theory of the vector equation (1) may be developed along the lines followed by Fredholm and Plemelj*.

§ 17. We have seen that in general the solution (12) of the non-homogeneous equation (1) becomes infinite at a singular value $\lambda_0$ of the parameter. In order that this should not be the case, i.e. in order that the equation (1) may admit a finite and continuous solution for this value of $\lambda$, certain conditions must be satisfied by the vector function $\mathbf{f}(t)$. To each of the $q$ linearly independent solutions (38) of the associated homogeneous equation will correspond one condition. Assume that (1) admits a solution $\mathbf{u}(t)$ for the parameter value $\lambda_0$, so that

$$\mathbf{u}(t) - \lambda_0 \int \mathbf{K}(ts) \bullet \mathbf{u}(s)\, ds = \mathbf{f}(t) \quad \ldots\ldots(39),$$

while simultaneously we have

$$\mathbf{v}_i(t) - \lambda_0 \int \mathbf{v}_i(\vartheta) \bullet \mathbf{K}(\vartheta t)\, d\vartheta = 0 \quad \ldots\ldots(40).$$

* "Zur Theorie der Fredholmschen Funktionalgleichung," *Monatshefte für Math. und Physik*, Bd. xv. (1904). Also *Potentialtheoretische Untersuchungen*, Teubner, Leipzig, 1911, S. 29–39.

Multiplying (39) by $\mathbf{v}_i(t)\cdot$ and integrating over the region $S$ we have

$$\int \mathbf{f}(t)\cdot\mathbf{v}_i(t)\,dt = \int \mathbf{u}(t)\cdot\mathbf{v}_i(t)\,dt - \lambda_0 \iint \mathbf{v}_i(t)\cdot\mathbf{K}(ts)\cdot\mathbf{u}(s)\,ds\,dt.$$

Now the direct product of any number of dyadics with a vector factor at either end or at both ends obeys the associative laws*: hence the order of integration may be changed in the last term as explained in § 5. The second member is therefore equal to

$$\int [\mathbf{v}_i(s) - \lambda_0 \int \mathbf{v}_i(t)\cdot\mathbf{K}(ts)\,dt]\cdot\mathbf{u}(s)\,ds$$

which vanishes in virtue of (40). The first member is therefore equal to zero, that is

$$\int \mathbf{f}(t)\cdot\mathbf{v}_i(t)\,dt = 0, \quad (i = 1, 2, \ldots, q) \qquad \ldots\ldots(41).$$

When two vector functions are such that the integral of their scalar product over a given region vanishes we shall say that the functions are *orthogonal* for that region. Hence we have proved that the necessary condition for the existence of a solution to the non-homogeneous equation (1) for a singular parameter value, is that the function $\mathbf{f}(t)$ be orthogonal to each of the $q$ linearly independent solutions of the associated homogeneous equation for that characteristic number. That this condition is also sufficient may be established by argument along the lines of the corresponding proof for the scalar integral equation†.

The conditions just stated may be applied to the systems of equations (2) and (4), for the vector equation (1) is equivalent to the system (2), and the homogeneous equation (15) to the system (4) with second members zero. It therefore follows that the necessary and sufficient condition that the system (2) may admit a finite and continuous set of solutions for a singular value of $\lambda$ is that the relation

$$\int [f_1(t)\,v_1(t) + f_2(t)\,v_2(t) + f_3(t)\,v_3(t)]\,dt = 0$$

should hold for every set of solutions $v_1(t)$, $v_2(t)$, $v_3(t)$ of the system (4) with second members zero. It is a *common mistake* to state this as though $v_1(t)$, $v_2(t)$, $v_3(t)$ were a set of solutions of the homogeneous system derived from (2) by interchanging the variables and putting the second members zero. This is quite wrong except in the particular case in which $K_{mn}(ts) = K_{nm}(ts)$ $(m, n = 1, 2, 3)$; that is when the dyadic $\mathbf{K}(ts)$ is self-conjugate.

§ 18. Two solutions $\mathbf{u}(t)$ and $\mathbf{v}(t)$ of the associated homogeneous integral equations

$$\mathbf{u}(t) = \lambda_1 \int \mathbf{K}(ts)\cdot\mathbf{u}(s)\,ds \qquad \ldots\ldots(42),$$

and

$$\mathbf{v}(t) = \lambda_2 \int \mathbf{v}(s)\cdot\mathbf{K}(st)\,ds \qquad \ldots\ldots(43),$$

corresponding to different characteristic numbers $\lambda_1$ and $\lambda_2$ are orthogonal to each other. For on multiplying the first by $\lambda_2\mathbf{v}(t)\cdot$ and integrating, the second by $\cdot\,\mathbf{u}(t)\lambda_1$ and integrating, and then subtracting the results we have

$$(\lambda_2 - \lambda_1)\int \mathbf{u}(t)\cdot\mathbf{v}(t)\,dt = \lambda_1\lambda_2 \iint \mathbf{v}(t)\cdot\mathbf{K}(ts)\cdot\mathbf{u}(s)\,ds\,dt - \lambda_1\lambda_2 \iint \mathbf{v}(s)\cdot\mathbf{K}(st)\cdot\mathbf{u}(t)\,ds\,dt.$$

But since the factors of the products of the second member are associative and therefore, as explained in the previous section, the order of integration may be changed, it follows that the second member is zero. Hence, because $\lambda_1$ and $\lambda_2$ are different, the integral of the left-hand side must vanish, giving

$$\int \mathbf{u}(t)\cdot\mathbf{v}(t)\,dt = 0.$$

The two solutions are therefore orthogonal.

* Gibbs-Wilson, *loc. cit.* p. 279.

† Cf. Fredholm, *loc. cit.* pp. 376–8; Plemelj, *Potent. Unter.* S. 36–7. Another proof will be found in § 24 below.

## VIII. The Conjugo-symmetric Dyadic Kernel.

§ 19. The question naturally arises whether there is a dyadic kernel playing for the vector equation (1) the same part as the symmetric kernel in the theory of the single scalar integral equation. Of such a dyadic kernel all that is required is that it make the homogeneous equation

$$\mathbf{u}(t) = \lambda \int \mathbf{K}(ts) \cdot \mathbf{u}(s)\, ds \quad \text{.................................(14)}$$

identical with its associated equation

$$\mathbf{v}(t) = \lambda \int \mathbf{v}(s) \cdot \mathbf{K}(st)\, ds \quad \text{.................................(15).}$$

The dyadic under consideration must therefore be such that $\mathbf{K}(ts)$ used as a prefactor is equivalent to $\mathbf{K}(st)$ used as a postfactor. A kernel presenting this property will be called *conjugo-symmetric.*

A conjugo-symmetric kernel need not be both self-conjugate as a dyadic and symmetric in the variables. In order that the two equations (14) and (15) should be identical the systems (2) and (4) must be identical: and vice versa. All that is necessary for this is that

$$K_{mn}(ts) = K_{nm}(st) \quad \text{.......................................(44),}$$
$$n, m = 1, 2, 3.$$

This property, if it exists, ensures conjugo-symmetry. It does not necessitate the symmetry of $K_{nm}(ts)$ in $t$ and $s$ unless $n$ and $m$ are equal: but $K_{11}(ts)$, $K_{22}(ts)$ and $K_{33}(ts)$ must be symmetric. Hence the scalar of a conjugo-symmetric dyadic kernel is symmetric in the variables. In order, however, that a kernel should be both self-conjugate and symmetric the relations (44) must hold, and in addition all these functions must be symmetric.

All the kernels formed by iteration from a conjugo-symmetric kernel are also conjugo-symmetric. For if $\mathbf{a}$ is any vector we have

$$\begin{aligned}\mathbf{K}_1(ts) \cdot \mathbf{a} &= \int \mathbf{K}(t\vartheta) \cdot \mathbf{K}(\vartheta s) \cdot \mathbf{a}\, d\vartheta \\ &= \int \mathbf{K}(t\vartheta) \cdot [\mathbf{a} \cdot \mathbf{K}(s\vartheta)]\, d\vartheta \\ &= \int \mathbf{a} \cdot \mathbf{K}(s\vartheta) \cdot \mathbf{K}(\vartheta t)\, d\vartheta \\ &= \mathbf{a} \cdot \mathbf{K}_1(st),\end{aligned}$$

each step following from the conjugo-symmetry of $\mathbf{K}(ts)$. The statement therefore holds for the first iterated kernel and may in exactly the same manner be established by induction for the $n$th, the first factor of the integrand in the above being replaced by $\mathbf{K}_{n-1}(t\vartheta)$.

§ 20. From the fundamental property that for a conjugo-symmetric kernel the equations (14) and (15) are identical may be deduced a set of important theorems corresponding to those that hold for the scalar integral equation with symmetric kernel.

(1) *Two solutions* $\mathbf{u}_1(t)$ *and* $\mathbf{u}_2(t)$ *of a homogeneous conjugo-symmetric vector integral equation corresponding to different characteristic numbers are orthogonal.*

This follows immediately from § 18, since the homogeneous equation is identical with its associated equation.

(2) *A real conjugo-symmetric kernel cannot have imaginary characteristic numbers.*

For suppose that $\alpha + i\beta$ is an imaginary root of $D(\lambda)$. Then $\alpha - i\beta$ is also a root.

Corresponding to the former is at least one solution of (14), which must be imaginary and therefore expressible in the form

$$\mathbf{u}_1(t) = \mathbf{p}(t) + i\mathbf{q}(t).$$

Hence, corresponding to the singular value $\alpha - i\beta$, there is the solution

$$\mathbf{u}_2(t) = \mathbf{p}(t) - i\mathbf{q}(t).$$

But by the last theorem $\mathbf{u}_1(t)$ and $\mathbf{u}_2(t)$ are orthogonal, so that

$$\int [\mathbf{p}(t) + i\mathbf{q}(t)] \cdot [\mathbf{p}(t) - i\mathbf{q}(t)]\, dt = 0.$$

The integrand is the scalar product of conjugate bivectors*, giving

$$\int [\mathbf{p}(t) \cdot \mathbf{p}(t) + \mathbf{q}(t) \cdot \mathbf{q}(t)]\, dt = 0;$$

that is, the sum of two positive quantities is zero, which is absurd. Hence the kernel does not admit an imaginary characteristic number.

(3) *Every conjugo-symmetric kernel possesses at least one characteristic number.*

The proof of this may be established along the lines of the proof for the scalar integral equation†.

(4) *For any characteristic number* $\lambda_0$ *the pole of the resolvent is simple.*

To prove this, take the relations (26 *a*) and (27 *b*). In the former interchange the letters $t$ and $s$. Multiply the equations by $\cdot \mathbf{B}_{r-1}(ts)$ and $\mathbf{B}_r(st) \cdot$ respectively and integrate with respect to $t$. On subtraction we have

$$\begin{aligned} 0 = \lambda_0 \iint & \mathbf{B}_r(s\vartheta) \cdot \mathbf{K}(\vartheta t) \cdot \mathbf{B}_{r-1}(ts)\, d\vartheta\, dt \\ - \lambda_0 \iint & \mathbf{B}_r(st) \cdot \mathbf{K}(t\vartheta) \cdot \mathbf{B}_{r-1}(\vartheta s)\, d\vartheta\, dt \\ + \iint & \mathbf{B}_r(st) \cdot \mathbf{K}(t\vartheta) \cdot \mathbf{B}_r(\vartheta s)\, d\vartheta\, dt. \end{aligned}$$

Since the order of integration may be changed, as previously shewn, the first two integrals are equal in magnitude and opposite in sign. The last integral may be simplified in virtue of (26 *a*) or (27 *b*), giving

$$\int \mathbf{B}_r(st) \cdot \mathbf{B}_r(ts)\, dt = 0.$$

Hence if $\mathbf{a}$ is any finite vector,

$$\int [\mathbf{a} \cdot \mathbf{B}_r(st)] \cdot [\mathbf{B}_r(ts) \cdot \mathbf{a}]\, dt = 0.$$

Now the dyadic $\mathbf{B}_r(ts)$ is conjugo-symmetric. For $\mathbf{B}_r(ts) \cdot \mathbf{a}$ and $\mathbf{a} \cdot \mathbf{B}_r(st)$ as functions of $t$ are simultaneous solutions of the associated equations (14) and (15) for the singular value $\lambda_0$: and as the equations are now identical these solutions are equal. The last relation may then be written

$$\int [\mathbf{a} \cdot \mathbf{B}_r(st)] \cdot [\mathbf{a} \cdot \mathbf{B}_r(st)]\, dt = 0,$$

that is, the integral of an essentially positive function vanishes, which necessitates the vanishing of the dyadic $\mathbf{B}_r(st)$ identically. This function therefore vanishes for all values of $r > 1$. Hence the pole of the resolvent is simple.

§ 21. *Principal system of fundamental functions.* The $q$ linearly independent solutions of the homogeneous equation (14) corresponding to a singular value $\lambda_1$ may be replaced by a system of $q$ normalised orthogonal‡ functions, that is, functions satisfying the relations

$$\int \mathbf{u}_r(t) \cdot \mathbf{u}_s(t)\, dt = \begin{matrix} 1 & \text{if} & r = s \\ 0 & \text{if} & r \neq s \end{matrix} \quad \ldots\ldots\ldots\ldots (45).$$

* Gibbs-Wilson, *loc. cit.* pp. 426–436. † Cf. Schmidt, *Math. Ann.* Bd. LXIII. (1907), S. 455–7. ‡ *Ibid.* S. 442–4.

Then since the fundamental functions* corresponding to different characteristic numbers are in the case of the conjugo-symmetric kernel orthogonal to each other, we may, when the kernel is of this nature, replace all the linearly independent solutions of the homogeneous equation corresponding to all the characteristic numbers by a series of normalised orthogonal fundamental functions satisfying the above relations. Such a system,

$$\mathbf{u}_1(t),\ \mathbf{u}_2(t),\ \ldots,\ \mathbf{u}_n(t) \qquad\qquad (46),$$

may be called the principal system of fundamental functions. We may assume that these functions are placed in the order of the increasing magnitude of the characteristic numbers to which they belong. Any solution of the homogeneous equation (14) is linearly expressible in terms of a finite number of functions of this series belonging to the same characteristic number.

§ 22. It may be possible to express the dyadic kernel $\mathbf{K}(ts)$ as the sum of a series of dyads whose antecedents and consequents are the vectors $\mathbf{u}_r(t)$, $\mathbf{u}_i(s)$ of the system (46).

Suppose it is possible to do this in such a way that the antecedents of the dyads are the successive functions (46), giving

$$\mathbf{K}(ts) = \mathbf{u}_1(t)\,\mathbf{c}_1 + \mathbf{u}_2(t)\,\mathbf{c}_2 + \ldots + \mathbf{u}_n(t)\,\mathbf{c}_n + \ldots \qquad\qquad (47),$$

where the consequents $\mathbf{c}_n$ are to be determined. If this relation holds, so that the second member is absolutely and uniformly convergent† when the number of terms is infinite, we may act on any vector $\mathbf{a}$ with each side as a prefactor, obtaining

$$\mathbf{K}(ts)\cdot\mathbf{a} = \mathbf{u}_1(t)\,\mathbf{c}_1\cdot\mathbf{a} + \mathbf{u}_2(t)\,\mathbf{c}_2\cdot\mathbf{a} + \ldots,$$

where all the terms are vectors. Multiplying each side by $\mathbf{u}_n(t)\cdot$ and integrating, we have, in virtue of the orthogonal relations (45),

$$[\int \mathbf{u}_n(t)\cdot\mathbf{K}(ts)\,dt]\cdot\mathbf{a} = \mathbf{c}_n\cdot\mathbf{a}.$$

This holds for any vector $\mathbf{a}$. It follows then that

$$\mathbf{c}_n = \int \mathbf{u}_n(t)\cdot\mathbf{K}(ts)\,dt = \frac{\mathbf{u}_n(s)}{\lambda_n}.$$

When, therefore, the representation (47) is possible, it becomes

$$\mathbf{K}(ts) = \frac{\mathbf{u}_1(t)\,\mathbf{u}_1(s)}{\lambda_1} + \frac{\mathbf{u}_2(t)\,\mathbf{u}_2(s)}{\lambda_2} + \ldots \qquad\qquad (48),$$

the functions $\mathbf{u}_n(t)$ forming the antecedents, and the functions $\mathbf{u}_n(s)$ the consequents of the dyads in the second member. This series corresponds to the bilinear series for the symmetric kernel of the scalar integral equation.

Conversely it may be shewn that if the series (48) is absolutely and uniformly convergent its sum is equal to the kernel $\mathbf{K}(ts)$.

§ 23. The following theorems for the conjugo-symmetric kernel may also be established without difficulty:

* The solutions of the homogeneous equation (14).

† That is to say, the series obtained by operating term by term on a finite vector is absolutely and uniformly convergent.

(1) *The necessary and sufficient condition that a continuous vector function* $\mathbf{h}(t)$ *may satisfy the identity*

$$\int \mathbf{K}(ts) \bullet \mathbf{h}(s)\, ds = 0 \quad \text{.....................................(49)}$$

*is that for all values of* $n$

$$\int \mathbf{u}_n(s) \bullet \mathbf{h}(s)\, ds = 0.$$

(2) *Any vector function* $\mathbf{g}(t)$ *expressible in the form*

$$\mathbf{g}(t) = \int \mathbf{K}(ts) \bullet \mathbf{l}(s)\, ds,$$

*where* $\mathbf{l}(s)$ *is a continuous vector function of the position of* $s$, *can be expanded in terms of the fundamental functions* (46) *according to the Fourier rule,*

$$\mathbf{g}(t) = \sum_n \mathbf{u}_n(t) \int \mathbf{g}(t) \bullet \mathbf{u}_n(t)\, dt$$

$$= \sum_n \frac{1}{\lambda_n} \mathbf{u}_n(t) \int \mathbf{l}(t) \bullet \mathbf{u}_n(t)\, dt.$$

We shall say that a conjugo-symmetric kernel is *closed* when there does not exist any continuous function $\mathbf{h}(t)$ satisfying the relation (49). It may be proved that

(3) *Every closed conjugo-symmetric kernel has an infinite number of singular parameter values.*

In conclusion, Schmidt's formula* for the solution of the non-homogeneous symmetric integral equation may be extended to the vector integral equation (1) where the kernel is conjugo-symmetric. Thus

(4) *If* $\lambda$ *is not equal to a characteristic number the equation* (1) *with conjugo-symmetric kernel has a unique continuous solution given by*

$$\mathbf{u}(t) = \mathbf{f}(t) - \lambda \sum_n \frac{\mathbf{u}_n(t)}{\lambda - \lambda_n} \int \mathbf{f}(t) \bullet \mathbf{u}_n(t)\, dt \quad \text{..........................(50).}$$

If however $\lambda$ is equal to a characteristic number $\lambda_m$ to which correspond $q$ fundamental functions of the principal system, for the above solution to remain finite it is necessary and sufficient that $\mathbf{f}(t)$ be orthogonal to each of these $q$ functions. If this condition is satisfied the solution of the integral equation becomes

$$\mathbf{u}(t) = \mathbf{f}(t) - \sum_1^q {}_i a_i \mathbf{u}_{m-i}(t) - \lambda_m \sum_n \frac{\mathbf{u}_n(t)}{\lambda_m - \lambda_n} \int \mathbf{f}(t) \bullet \mathbf{u}_n(t)\, dt \quad \text{..............(51),}$$

where the quantities $a_i$ are arbitrary constants, the first summation including the fundamental functions belonging to $\lambda_m$, and the second the remaining ones of the principal system (46). That $\mathbf{u}(t)$ given by (51) is a solution of (1) for the singular parameter value $\lambda_m$ is easily verified by direct substitution.

§ 24. From the results established above for the conjugo-symmetric kernel a proof may be deduced† of the sufficiency of the conditions found in § 17 for the existence of a solution to the non-homogeneous equation

$$\mathbf{u}(t) - \lambda_0 \int \mathbf{K}(ts) \bullet \mathbf{u}(s)\, ds = \mathbf{f}(t) \quad \text{.............................(39),}$$

* Cf. Schmidt, *loc. cit.* S. 454. † *Ibid.* § 13.

where $\lambda_0$ is a characteristic number of the kernel, not supposed symmetric or conjugo-symmetric. From this equation and its associated equation

$$\mathbf{v}(t) - \lambda_0 \int \mathbf{v}(s) \cdot \mathbf{K}(st)\, ds = \mathbf{g}(t) \quad \text{.......}(52),$$

we readily deduce

$$\mathbf{g}(t) - \lambda_0 \int \mathbf{K}(ts) \cdot \mathbf{g}(s)\, ds = \mathbf{v}(t) - \lambda_0 \int \mathbf{Q}(ts) \cdot \mathbf{v}(s)\, ds \quad \text{.......}(53),$$

and
$$\int \mathbf{g}(t) \cdot \mathbf{g}(t)\, dt = \int \mathbf{v}(t) \cdot [\mathbf{v}(t) - \lambda_0 \int \mathbf{Q}(ts) \cdot \mathbf{v}(s)\, ds]\, dt \quad \text{.......}(54),$$

where
$$\mathbf{Q}(st) = \mathbf{K}(st) + \mathbf{K}_c(ts) - \lambda_0 \int \mathbf{K}(s\sigma) \cdot \mathbf{K}_c(t\sigma)\, d\sigma\,;$$

$\mathbf{K}_c(ts)$ denoting the conjugate dyadic of $\mathbf{K}(ts)$, that is the dyadic which acting as a prefactor is equivalent to $\mathbf{K}(ts)$ used as a postfactor, and conversely. The kernel $\mathbf{Q}(st)$ is conjugo-symmetric. For if $\mathbf{r}$ is any finite vector

$$\begin{aligned}\mathbf{Q}(st) \cdot \mathbf{r} &= \mathbf{K}(st) \cdot \mathbf{r} + \mathbf{K}_c(ts) \cdot \mathbf{r} - \lambda_0 \int \mathbf{K}(s\sigma) \cdot \mathbf{K}_c(t\sigma) \cdot \mathbf{r}\, d\sigma \\ &= \mathbf{r} \cdot \mathbf{K}_c(st) + \mathbf{r} \cdot \mathbf{K}(ts) - \lambda_0 \int \mathbf{r} \cdot \mathbf{K}(t\sigma) \cdot \mathbf{K}_c(s\sigma)\, d\sigma.\end{aligned}$$

The transformation of the integral is a consequence of the theorem that the conjugate of the product of two dyadics is equal to the product of their conjugates taken in the reverse order*. The last equation may be written

$$\mathbf{Q}(st) \cdot \mathbf{r} = \mathbf{r} \cdot \mathbf{Q}(ts),$$

proving that $\mathbf{Q}(st)$ is conjugo-symmetric.

Now if $\mathbf{v}(t)$ is a solution of the associated homogeneous equation

$$\mathbf{v}(t) - \lambda_0 \int \mathbf{v}(s) \cdot \mathbf{K}(st)\, ds = 0 \quad \text{.......}(40)$$

obtained from (52) by putting $\mathbf{g}(t)$ zero, it follows from (53) that it is also a fundamental function of $\mathbf{Q}(st)$ for the characteristic number $\lambda_0$. Conversely, if $\mathbf{v}(t)$ is a fundamental function of $\mathbf{Q}(st)$ for this parameter value, it follows from (54) and (52) that it is also a solution of (40). The linearly independent solutions of (40) are therefore identical with the fundamental functions of $\mathbf{Q}(st)$.

If now we transform (39) by the substitution

$$\mathbf{u}(t) = \mathbf{w}(t) - \lambda_0 \int \mathbf{w}(s) \cdot \mathbf{K}(st)\, ds \quad \text{.......}(55)$$

it becomes
$$\mathbf{w}(t) - \lambda_0 \int \mathbf{Q}(ts) \cdot \mathbf{w}(s)\, ds = \mathbf{f}(t).$$

But by the preceding section the necessary and sufficient condition that this equation may admit a finite solution is the orthogonality of $\mathbf{f}(t)$ to all the fundamental functions of $\mathbf{Q}(st)$ corresponding to $\lambda_0$, that is to all the linearly independent solutions of (40). Therefore in virtue of the relation (55) this is the necessary and sufficient condition that (39) may admit a solution for the singular parameter value $\lambda_0$.

§ 25. *Singular Kernel.* We have up to the present assumed that the dyadic $\mathbf{K}(ts)$ remains finite, i.e. that all the coefficients $K_{ir}(ts)$ $[i, r = 1, 2, 3]$ in its nonion form are finite for every point $t$, $s$ of the region $S$. In many physical problems when the points $t$ and $s$ coalesce the kernel $\mathbf{K}(ts)$ becomes infinite like $1/r^\alpha$, $r$ being the distance between the two points. But as shewn in § 6 the integral equation (1) may be replaced by (8) in which the kernel is the $n$th iterated kernel of $\mathbf{K}(ts)$. When $\alpha$ is not too large, this iterated kernel will be everywhere finite if $n$ is sufficiently great, and the methods of the

* Gibbs-Wilson, *loc. cit.* p. 294.

preceding pages will be applicable. It may be shewn, as in the theory of the scalar integral equation*, that if the region $S$ of integration is a given surface this method holds for $\alpha < 2$, and if a given volume for $\alpha < 3$.

The theorems of §§ 19–23 on the conjugo-symmetric kernel are true when $\mathbf{K}(ts)$ is singular, provided its discontinuities are regularly distributed and the integrals

$$\int [\mathbf{r} \cdot \mathbf{K}(st)]^2 ds, \quad \text{and} \quad \int [\mathbf{K}(ts) \cdot \mathbf{r}]^2 ds \dagger$$

are finite and continuous functions of the position of $t$, $\mathbf{r}$ being any finite and continuous vector.

In the foregoing pages the author has dealt only with the theory of the vector integral equation. Applications of this theory to various problems of mathematical physics will be discussed elsewhere.

* Cf. Fredholm, *loc. cit.* pp. 384–390; Heywood and Fréchet, *L'équation de Fredholm &c.*, Paris (1912), pp. 141–145.

† The square of a vector denotes as usual the scalar product of the vector by itself. This is equal to the square of its tensor.

# IX. *On certain Arithmetical Functions.*

By S. Ramanujan.

[Communicated by G. H. Hardy*.]

[*Received and Read* 25 October 1915.]

1. Let $\sigma_s(n)$ denote the sum of the $s$th powers of the divisors of $n$ (including 1 and $n$), and let

$$\sigma_s(0) = \tfrac{1}{2}\zeta(-s),$$

where $\zeta(s)$ is the Riemann Zeta-function. Further let

$$\Sigma_{r,s}(n) = \sigma_r(0)\,\sigma_s(n) + \sigma_r(1)\,\sigma_s(n-1) + \ldots + \sigma_r(n)\,\sigma_s(0) \quad \ldots\ldots\ldots\ldots\ldots(1).$$

In this paper I prove that

$$\Sigma_{r,s}(n) = \frac{\Gamma(r+1)\,\Gamma(s+1)}{\Gamma(r+s+2)}\frac{\zeta(r+1)\,\zeta(s+1)}{\zeta(r+s+2)}\sigma_{r+s+1}(n)$$
$$+\frac{\zeta(1-r)+\zeta(1-s)}{r+s}n\sigma_{r+s-1}(n) + O\{n^{\frac{2}{3}(r+s+1)}\} \quad \ldots\ldots\ldots\ldots(2),$$

whenever $r$ and $s$ are positive odd integers. I also prove that there is no error term on the right-hand side of (2) in the following nine cases: $r=1, s=1$; $r=1, s=3$; $r=1, s=5$; $r=1, s=7$; $r=1, s=11$; $r=3, s=3$; $r=3, s=5$; $r=3, s=9$; $r=5, s=7$. That is to say $\Sigma_{r,s}(n)$ has a finite expression in terms of $\sigma_{r+s+1}(n)$ and $\sigma_{r+s-1}(n)$ in these nine cases; but for other values of $r$ and $s$ it involves other arithmetical functions as well.

It appears probable, from the empirical results I obtain in §§ 18—23, that the error term on the right-hand side of (2) is of the form

$$O\{n^{\frac{1}{2}(r+s+1+\epsilon)}\} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots(3),$$

where $\epsilon$ is any positive number, and not of the form

$$o\{n^{\frac{1}{2}(r+s+1)}\} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots(4).$$

But all I can prove rigorously is (i) that the error is of the form

$$O\{n^{\frac{2}{3}(r+s+1)}\}$$

in all cases, (ii) that it is of the form

$$O\{n^{\frac{2}{3}(r+s+\frac{4}{5})}\} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots(5)$$

if $r+s$ is of the form $6m$, (iii) that it is of the form

$$O\{n^{\frac{2}{3}(r+s+\frac{1}{2})}\} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots(6)$$

if $r+s$ is of the form $6m+4$, and (iv) that it is not of the form

$$o\{n^{\frac{1}{2}(r+s)}\} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots(7).$$

* I am indebted to Mr Hardy for his kind assistance and advice.

It follows from (2) that, if $r$ and $s$ are positive odd integers, then

$$\Sigma_{r,s}(n) \sim \frac{\Gamma(r+1)\Gamma(s+1)}{\Gamma(r+s+2)} \frac{\zeta(r+1)\zeta(s+1)}{\zeta(r+s+2)} \sigma_{r+s+1}(n) \quad \text{.................(8).}$$

It seems very likely that (8) is true for all positive values of $r$ and $s$, but this I am at present unable to prove.

2. If $\Sigma_{r,s}(n)/\sigma_{r+s+1}(n)$ tends to a limit, then the limit must be

$$\frac{\Gamma(r+1)\Gamma(s+1)}{\Gamma(r+s+2)} \frac{\zeta(r+1)\zeta(s+1)}{\zeta(r+s+2)}.$$

For then

$$\lim_{n\to\infty} \frac{\Sigma_{r,s}(n)}{\sigma_{r+s+1}(n)} = \lim_{n\to\infty} \frac{\Sigma_{r,s}(1)+\Sigma_{r,s}(2)+\ldots+\Sigma_{r,s}(n)}{\sigma_{r+s+1}(1)+\sigma_{r+s+1}(2)+\ldots+\sigma_{r+s+1}(n)}$$

$$= \lim_{x\to 1} \frac{\Sigma_{r,s}(0)+\Sigma_{r,s}(1)x+\Sigma_{r,s}(2)x^2+\ldots}{\sigma_{r+s+1}(0)+\sigma_{r+s+1}(1)x+\sigma_{r+s+1}(2)x^2+\ldots}$$

$$= \lim_{x\to 1} \frac{S_r S_s}{S_{r+s+1}},$$

where

$$S_r = \tfrac{1}{2}\zeta(-r) + \frac{1^r x}{1-x} + \frac{2^r x^2}{1-x^2} + \frac{3^r x^3}{1-x^3} + \ldots \quad \text{.........................(9).}$$

Now it is known that, if $r > 0$, then

$$S_r \sim \frac{\Gamma(r+1)\zeta(r+1)}{(1-x)^{r+1}} \quad \text{...................................(10),}$$

as $x \to 1$*. Hence we obtain the result stated.

3. It is easy to see that

$$\sigma_r(1)+\sigma_r(2)+\sigma_r(3)+\ldots+\sigma_r(n)$$
$$= u_1+u_2+u_3+u_4+\ldots+u_n,$$

where

$$u_t = 1^r + 2^r + 3^r + \ldots + \left[\frac{n}{t}\right]^r.$$

From this it is easy to deduce that

$$\sigma_r(1)+\sigma_r(2)+\ldots+\sigma_r(n) \sim \frac{n^{r+1}}{r+1}\zeta(r+1) \quad \text{.....................(11)†}$$

and

$$\sigma_r(1)(n-1)^s + \sigma_r(2)(n-2)^s + \ldots + \sigma_r(n-1)1^s \sim \frac{\Gamma(r+1)\Gamma(s+1)}{\Gamma(r+s+2)}\zeta(r+1)n^{r+s+1},$$

provided $r > 0$, $s \geqslant 0$. Now

$$\sigma_s(n) > n^s,$$

and

$$\sigma_s(n) < n^s(1^{-s}+2^{-s}+3^{-s}+\ldots) = n^s\zeta(s).$$

From these inequalities and (1) it follows that

$$\underline{\lim} \frac{\Sigma_{r,s}(n)}{n^{r+s+1}} \geqslant \frac{\Gamma(r+1)\Gamma(s+1)}{\Gamma(r+s+2)}\zeta(r+1) \quad \text{.............................(12),}$$

if $r>0$ and $s \geqslant 0$; and

$$\overline{\lim} \frac{\Sigma_{r,s}(n)}{n^{r+s+1}} \leqslant \frac{\Gamma(r+1)\Gamma(s+1)}{\Gamma(r+s+2)}\zeta(r+1)\zeta(s) \quad \text{.......................(13),}$$

* Knopp, *Dissertation* (Berlin, 1907), p. 34.

† (10) follows from this as an immediate corollary.

if $r>0$ and $s>1$. Thus $n^{-r-s-1}\Sigma_{r,s}(n)$ oscillates between limits included in the interval

$$\frac{\Gamma(r+1)\Gamma(s+1)}{\Gamma(r+s+2)}\zeta(r+1), \qquad \frac{\Gamma(r+1)\Gamma(s+1)}{\Gamma(r+s+2)}\zeta(r+1)\zeta(s).$$

On the other hand $n^{-r-s-1}\sigma_{r+s+1}(n)$ oscillates between 1 and $\zeta(r+s+1)$, assuming values as near as we please to either of these limits. The formula (8) shows that the actual limits of indetermination of $n^{-r-s-1}\Sigma_{r,s}(n)$ are

$$\frac{\Gamma(r+1)\Gamma(s+1)}{\Gamma(r+s+2)}\frac{\zeta(r+1)\zeta(s+1)}{\zeta(r+s+2)}, \qquad \frac{\Gamma(r+1)\Gamma(s+1)}{\Gamma(r+s+2)}\frac{\zeta(r+1)\zeta(s+1)\zeta(r+s+1)}{\zeta(r+s+2)} \quad \text{.........(14).}$$

Naturally

$$\zeta(r+1) < \frac{\zeta(r+1)\zeta(s+1)}{\zeta(r+s+2)} < \frac{\zeta(r+1)\zeta(s+1)\zeta(r+s+1)}{\zeta(r+s+2)} < \zeta(r+1)\zeta(s)^*.$$

What is remarkable about the formula (8) is that it shows the asymptotic equality of two functions neither of which itself increases in a regular manner.

4. It is easy to see that, if $n$ is a positive integer, then

$$\cot\tfrac{1}{2}\theta\sin n\theta = 1 + 2\cos\theta + 2\cos 2\theta + \ldots + 2\cos(n-1)\theta + \cos n\theta.$$

Suppose now that

$$\left(\frac{1}{4}\cot\frac{1}{2}\theta + \frac{x\sin\theta}{1-x} + \frac{x^2\sin 2\theta}{1-x^2} + \frac{x^3\sin 3\theta}{1-x^3} + \ldots\right)^2$$
$$= (\tfrac{1}{4}\cot\tfrac{1}{2}\theta)^2 + C_0 + C_1\cos\theta + C_2\cos 2\theta + C_3\cos 3\theta + \ldots,$$

where $C_n$ is independent of $\theta$. Then we have

$$\begin{aligned} C_0 &= \frac{1}{2}\left(\frac{x}{1-x} + \frac{x^2}{1-x^2} + \frac{x^3}{1-x^3} + \ldots\right) \\ &\quad + \frac{1}{2}\left\{\left(\frac{x}{1-x}\right)^2 + \left(\frac{x^2}{1-x^2}\right)^2 + \left(\frac{x^3}{1-x^3}\right)^2 + \ldots\right\} \\ &= \frac{1}{2}\left\{\frac{x}{(1-x)^2} + \frac{x^2}{(1-x^2)^2} + \frac{x^3}{(1-x^3)^2} + \ldots\right\} \\ &= \frac{1}{2}\left\{\frac{x}{1-x} + \frac{2x^2}{1-x^2} + \frac{3x^3}{1-x^3} + \ldots\right\} \quad \text{..............................(15).} \end{aligned}$$

Again

$$\begin{aligned} C_n &= \frac{1}{2}\left(\frac{x^n}{1-x^n} + \frac{x^{n+1}}{1-x^{n+1}} + \frac{x^{n+2}}{1-x^{n+2}} + \frac{x^{n+3}}{1-x^{n+3}} + \ldots\right) \\ &\quad + \frac{x}{1-x}\cdot\frac{x^{n+1}}{1-x^{n+1}} + \frac{x^2}{1-x^2}\cdot\frac{x^{n+2}}{1-x^{n+2}} + \frac{x^3}{1-x^3}\cdot\frac{x^{n+3}}{1-x^{n+3}} + \ldots \\ &\quad - \frac{1}{2}\left\{\frac{x}{1-x}\cdot\frac{x^{n-1}}{1-x^{n-1}} + \frac{x^2}{1-x^2}\cdot\frac{x^{n-2}}{1-x^{n-2}} + \ldots + \frac{x^{n-1}}{1-x^{n-1}}\cdot\frac{x}{1-x}\right\}. \end{aligned}$$

* For example when $r=1$ and $s=9$ this inequality becomes

$$1{\cdot}64493\ldots < 1{\cdot}64616\ldots < 1{\cdot}64697\ldots < 1{\cdot}64823\ldots.$$

Hence

$$\begin{aligned}\frac{C_n}{x^n}(1-x^n) &= \frac{1}{2} + \left(\frac{x}{1-x} - \frac{x^{n+1}}{1-x^{n+1}}\right) + \left(\frac{x^2}{1-x^2} - \frac{x^{n+2}}{1-x^{n+2}}\right) + \dots \\ &\quad - \frac{1}{2}\left\{\left(1 + \frac{x}{1-x} + \frac{x^{n-1}}{1-x^{n-1}}\right) + \left(1 + \frac{x^2}{1-x^2} + \frac{x^{n-2}}{1-x^{n-2}}\right) + \dots\right. \\ &\qquad\qquad \left. + \left(1 + \frac{x^{n-1}}{1-x^{n-1}} + \frac{x}{1-x}\right)\right\} \\ &= \frac{1}{1-x^n} - \frac{n}{2}.\end{aligned}$$

That is to say

$$C_n = \frac{x^n}{(1-x^n)^2} - \frac{nx^n}{2(1-x^n)} \qquad \text{(16)}.$$

It follows that

$$\begin{aligned}&\left(\frac{1}{4}\cot\frac{1}{2}\theta + \frac{x\sin\theta}{1-x} + \frac{x^2\sin 2\theta}{1-x^2} + \frac{x^3\sin 3\theta}{1-x^3} + \dots\right)^2 \\ &= \left(\frac{1}{4}\cot\frac{1}{2}\theta\right)^2 + \frac{x\cos\theta}{(1-x)^2} + \frac{x^2\cos 2\theta}{(1-x^2)^2} + \frac{x^3\cos 3\theta}{(1-x^3)^2} + \dots \\ &\quad + \frac{1}{2}\left\{\frac{x}{1-x}(1-\cos\theta) + \frac{2x^2}{1-x^2}(1-\cos 2\theta) + \frac{3x^3}{1-x^3}(1-\cos 3\theta) + \dots\right\} \qquad \text{(17)}.\end{aligned}$$

Similarly, using the equation

$$\cot^2\tfrac{1}{2}\theta\,(1-\cos n\theta) = (2n-1) + 4(n-1)\cos\theta + 4(n-2)\cos 2\theta + \dots + 4\cos(n-1)\theta + \cos n\theta,$$

we can show that

$$\begin{aligned}&\left\{\frac{1}{8}\cot^2\frac{1}{2}\theta + \frac{1}{12} + \frac{x}{1-x}(1-\cos\theta) + \frac{2x^2}{1-x^2}(1-\cos 2\theta) + \frac{3x^3}{1-x^3}(1-\cos 3\theta) + \dots\right\}^2 \\ &= \left(\frac{1}{8}\cot^2\frac{1}{2}\theta + \frac{1}{12}\right)^2 + \frac{1}{12}\left\{\frac{1^3x}{1-x}(5+\cos\theta) + \frac{2^3x^2}{1-x^2}(5+\cos 2\theta)\right. \\ &\qquad \left. + \frac{3^3x^3}{1-x^3}(5+\cos 3\theta) + \dots\right\} \qquad \text{(18)}.\end{aligned}$$

For example, putting $\theta = \frac{2}{3}\pi$ and $\theta = \frac{1}{2}\pi$ in (17), we obtain

$$\begin{aligned}&\left(\frac{1}{6} + \frac{x}{1-x} - \frac{x^2}{1-x^2} + \frac{x^4}{1-x^4} - \frac{x^5}{1-x^5} + \dots\right)^2 \\ &\qquad = \frac{1}{36} + \frac{1}{3}\left(\frac{x}{1-x} + \frac{2x^2}{1-x^2} + \frac{4x^4}{1-x^4} + \frac{5x^5}{1-x^5} + \dots\right) \qquad \text{(19)},\end{aligned}$$

where 1, 2, 4, 5, ... are the natural numbers without the multiples of 3; and

$$\begin{aligned}&\left(\frac{1}{4} + \frac{x}{1-x} - \frac{x^3}{1-x^3} + \frac{x^5}{1-x^5} - \frac{x^7}{1-x^7} + \dots\right)^2 \\ &\qquad = \frac{1}{16} + \frac{1}{2}\left(\frac{x}{1-x} + \frac{2x^2}{1-x^2} + \frac{3x^3}{1-x^3} + \frac{5x^5}{1-x^5} + \dots\right) \qquad \text{(20)},\end{aligned}$$

where 1, 2, 3, 5, ... are the natural numbers without the multiples of 4.

5. It follows from (18) that

$$\left(\frac{1}{2\theta^2}+\frac{\theta^2}{2!}S_3-\frac{\theta^4}{4!}S_5+\frac{\theta^6}{6!}S_7-\dots\right)^2$$
$$=\frac{1}{4\theta^4}+\frac{1}{2}S_3-\frac{1}{12}\left(\frac{\theta^2}{2!}S_5-\frac{\theta^4}{4!}S_7+\frac{\theta^6}{6!}S_9-\dots\right)\dots\dots\dots(21),$$

where $S_r$ is the same as in (9). Equating the coefficients of $\theta^n$ in both sides in (21), we obtain

$$\frac{(n-2)(n+5)}{12(n+1)(n+2)}S_{n+3}={}^nC_2S_3S_{n-1}+{}^nC_4S_5S_{n-3}+{}^nC_6S_7S_{n-5}+\dots+{}^nC_{n-2}S_{n-1}S_3\dots(22),$$

where
$${}^nC_r=\frac{n!}{r!(n-r)!},$$

if $n$ is an even integer greater than 2.

Let us now suppose that

$$\Phi_{r,s}(x)=\sum_{m=1}^{m=\infty}\sum_{n=1}^{n=\infty}m^rn^sx^{mn}\dots\dots\dots(23),$$

so that
$$\Phi_{r,s}(x)=\Phi_{s,r}(x),$$

and
$$\left.\begin{aligned}\Phi_{0,s}(x)&=\frac{1^sx}{1-x}+\frac{2^sx^2}{1-x^2}+\frac{3^sx^3}{1-x^3}+\dots=S_s-\tfrac{1}{2}\zeta(-s)\\ \Phi_{1,s}(x)&=\frac{1^sx}{(1-x)^2}+\frac{2^sx^2}{(1-x^2)^2}+\frac{3^sx^3}{(1-x^3)^2}+\dots\end{aligned}\right\}\dots\dots\dots(24).$$

Further let

$$\left.\begin{aligned}P&=-\ 24S_1=1-\ 24\left(\frac{x}{1-x}+\frac{2x^2}{1-x^2}+\frac{3x^3}{1-x^3}+\dots\right)^*\\ Q&=\ 240S_3=1+240\left(\frac{1^3x}{1-x}+\frac{2^3x^2}{1-x^2}+\frac{3^3x^3}{1-x^3}+\dots\right)\\ R&=-504S_5=1-504\left(\frac{1^5x}{1-x}+\frac{2^5x^2}{1-x^2}+\frac{3^5x^3}{1-x^3}+\dots\right)\end{aligned}\right\}\dots\dots\dots(25).$$

Then putting $n=4, 6, 8, \dots$ in (22) we obtain the results contained in the following table.

TABLE I.

1. $1-24\Phi_{0,1}(x)=P.$
2. $1+240\Phi_{0,3}(x)=Q.$
3. $1-504\Phi_{0,5}(x)=R.$
4. $1+480\Phi_{0,7}(x)=Q^2.$
5. $1-264\Phi_{0,9}(x)=QR.$

* If $x=q^2$, then in the notation of elliptic functions

$$P=\frac{12\eta\omega}{\pi^2}=\left(\frac{2K}{\pi}\right)^2\left(\frac{3E}{K}+k^2-2\right),$$

$$Q=\frac{12g_2\omega^4}{\pi^4}=\left(\frac{2K}{\pi}\right)^4(1-k^2+k^4),$$

$$R=\frac{216g_3\omega^6}{\pi^6}=\left(\frac{2K}{\pi}\right)^6(1+k^2)(1-2k^2)(1-\tfrac{1}{2}k^2).$$

TABLE I (*continued*).

6. $691 + 65520\Phi_{0,11}(x) = 441Q^3 + 250R^2.$
7. $1 - 24\Phi_{0,13}(x) = Q^2R.$
8. $3617 + 16320\Phi_{0,15}(x) = 1617Q^4 + 2000QR^2.$
9. $43867 - 28728\Phi_{0,17}(x) = 38367Q^3R + 5500R^3.$
10. $174611 + 13200\Phi_{0,19}(x) = 53361Q^5 + 121250Q^2R^2.$
11. $77683 - 552\Phi_{0,21}(x) = 57183Q^4R + 20500QR^3.$
12. $236364091 + 131040\Phi_{0,23}(x) = 49679091Q^6 + 176400000Q^3R^2 + 10285000R^4.$
13. $657931 - 24\Phi_{0,25}(x) = 392931Q^5R + 265000Q^2R^3.$
14. $3392780147 + 6960\Phi_{0,27}(x) = 489693897Q^7 + 2507636250Q^4R^2 + 395450000QR^4.$
15. $1723168255201 - 171864\Phi_{0,29}(x) = 815806500201Q^6R + 881340705000Q^3R^3 + 26021050000R^5.$
16. $7709321041217 + 32640\Phi_{0,31}(x) = 764412173217Q^8 + 5323905468000Q^5R^2 + 1621003400000Q^2R^4.$

In general $$\tfrac{1}{2}\zeta(-s) + \Phi_{0,s}(x) = \Sigma K_{m,n} Q^m R^n \qquad \text{(26)},$$
where $K_{m,n}$ is a constant and $m$ and $n$ are positive integers (including zero) satisfying the equation
$$4m + 6n = s + 1.$$
This is easily proved by induction, using (22).

6. Again from (17) we have
$$\left(\frac{1}{2\theta} + \frac{\theta}{1!}S_1 - \frac{\theta^3}{3!}S_3 + \frac{\theta^5}{5!}S_5 - \dots\right)^2$$
$$= \frac{1}{4\theta^2} + S_1 - \frac{\theta^2}{2!}\Phi_{1,2}(x) + \frac{\theta^4}{4!}\Phi_{1,4}(x) - \frac{\theta^6}{6!}\Phi_{1,6}(x) + \dots$$
$$+ \frac{1}{2}\left(\frac{\theta^2}{2!}S_3 - \frac{\theta^4}{4!}S_5 + \frac{\theta^6}{6!}S_7 - \right) \qquad \text{(27)}.$$

Equating the coefficients of $\theta^n$ in both sides in (27) we obtain
$$\frac{n+3}{2(n+1)}S_{n+1} - \Phi_{1,n}(x) = {}^nC_1S_1S_{n-1} + {}^nC_3S_3S_{n-3} + {}^nC_5S_5S_{n-5} + \dots + {}^nC_{n-1}S_{n-1}S_1 \qquad \text{(28)},$$
if $n$ is a positive even integer. From this we deduce the results contained in Table II.

TABLE II.

1. $288\Phi_{1,2}(x) = Q - P^2.$
2. $720\Phi_{1,4}(x) = PQ - R.$
3. $1008\Phi_{1,6}(x) = Q^2 - PR.$
4. $720\Phi_{1,8}(x) = Q(PQ - R).$
5. $1584\Phi_{1,10}(x) = 3Q^3 + 2R^2 - 5PQR.$
6. $65520\Phi_{1,12}(x) = P(441Q^3 + 250R^2) - 691Q^2R.$
7. $144\Phi_{1,14}(x) = Q(3Q^3 + 4R^2 - 7PQR).$

In general $$\Phi_{1,s}(x) = \Sigma K_{l,m,n} P^l Q^m R^n \qquad \text{(29)},$$
where $l \leqslant 2$ and $2l + 4m + 6n = s + 2$. This is easily proved by induction, using (28).

7. We have

$$\left.\begin{aligned} x\frac{dP}{dx} &= -\ \ 24\Phi_{1,2}(x) = \frac{P^2-Q}{12} \\ x\frac{dQ}{dx} &= \ \ \ 240\Phi_{1,4}(x) = \frac{PQ-R}{3} \\ x\frac{dR}{dx} &= -504\Phi_{1,6}(x) = \frac{PR-Q^2}{2} \end{aligned}\right\} \qquad \text{(30)}.$$

Suppose now that $r < s$ and that $r + s$ is even. Then

$$\Phi_{r,s}(x) = \left(x\frac{d}{dx}\right)^r \Phi_{0,s-r}(x) \qquad \text{(31)},$$

and $\Phi_{0,s-r}(x)$ is a polynomial in $Q$ and $R$. Also

$$x\frac{dP}{dx}, \quad x\frac{dQ}{dx}, \quad x\frac{dR}{dx}$$

are polynomials in $P$, $Q$ and $R$. Hence $\Phi_{r,s}(x)$ is a polynomial in $P$, $Q$ and $R$. Thus we deduce the results contained in Table III.

TABLE III.

1. $1728\Phi_{2,3}(x) = 3PQ - 2R - P^3.$
2. $1728\Phi_{2,5}(x) = P^2Q - 2PR + Q^2.$
3. $1728\Phi_{2,7}(x) = 2PQ^2 - P^2R - QR.$
4. $8640\Phi_{2,9}(x) = 9P^2Q^2 - 18PQR + 5Q^3 + 4R^2.$
5. $1728\Phi_{2,11}(x) = 6PQ^3 - 5P^2QR + 4PR^2 - 5Q^2R.$
6. $6912\Phi_{3,4}(x) = 6P^2Q - 8PR + 3Q^2 - P^4.$
7. $3456\Phi_{3,6}(x) = P^3Q - 3P^2R + 3PQ^2 - QR.$
8. $5184\Phi_{3,8}(x) = 6P^2Q^2 - 2P^3R - 6PQR + Q^3 + R^2.$
9. $20736\Phi_{4,5}(x) = 15PQ^2 - 20P^2R + 10P^3Q - 4QR - P^5.$
10. $41472\Phi_{4,7}(x) = 7(P^4Q - 4P^3R + 6P^2Q^2 - 4PQR) + 3Q^3 + 4R^2.$

In general

$$\Phi_{r,s}(x) = \Sigma K_{l,m,n} P^l Q^m R^n \qquad \text{(32)},$$

where $l - 1$ does not exceed the smaller of $r$ and $s$ and

$$2l + 4m + 6n = r + s + 1.$$

The results contained in these three tables are of course really results in the theory of elliptic functions. For example $Q$ and $R$ are substantially the invariants $g_2$ and $g_3$, and the formulae of Table I are equivalent to the formulae which express the coefficients in the series

$$\wp(u) = \frac{1}{u^2} + \frac{g_2 u^2}{20} + \frac{g_3 u^4}{28} + \frac{g_2^2 u^6}{1200} + \frac{3g_2 g_3 u^8}{6160} + \dots$$

in terms of $g_2$ and $g_3$. The elementary proof of these formulae given in the preceding sections seems to be of some interest in itself.

8. In what follows we shall require to know the form of $\Phi_{1,s}(x)$ more precisely than is shown by the formula (29).

We have
$$\tfrac{1}{2}\zeta(-s)+\Phi_{0,s}(x)=\Sigma K_{m,n}Q^m R^n \quad\text{.....(33)},$$
where $s$ is an odd integer greater than 1 and $4m+6n=s+1$. Also
$$x\frac{d}{dx}(Q^m R^n)=\left(\frac{m}{3}+\frac{n}{2}\right)PQ^m R^n-\left(\frac{m}{3}Q^{m-1}R^{n+1}+\frac{n}{3}Q^{m-2}R^{n-1}\right) \quad\text{.....(34)}.$$
Differentiating (33) and using (34) we obtain
$$\Phi_{1,s-1}(x)=\tfrac{1}{12}(s+1)P\{\tfrac{1}{2}\zeta(-s)+\Phi_{0,s}(x)\}+\Sigma K_{m,n}Q^m R^n \quad\text{.....(35)},$$
where $s$ is an odd integer greater than 1 and $4m+6n=s+3$. But when $s=1$ we have
$$\Phi_{1,2}(x)=\frac{Q-P^2}{288} \quad\text{.....(36)}.$$

9. Suppose now that
$$F_{r,s}(x)=\{\tfrac{1}{2}\zeta(-r)+\Phi_{0,r}(x)\}\{\tfrac{1}{2}\zeta(-s)+\Phi_{0,s}(x)\}$$
$$-\frac{\zeta(1-r)+\zeta(1-s)}{r+s}\Phi_{1,r+s}(x)-\frac{\Gamma(r+1)\Gamma(s+1)}{\Gamma(r+s+2)}\frac{\zeta(r+1)\zeta(s+1)}{\zeta(r+s+2)}$$
$$\times\{\tfrac{1}{2}\zeta(-r-s-1)+\Phi_{0,r+s+1}(x)\} \quad\text{.....(37)}.$$

Then it follows from (33), (35) and (36) that, if $r$ and $s$ are positive odd integers,
$$F_{r,s}(x)=\Sigma K_{m,n}Q^m R^n \quad\text{.....(38)},$$
where
$$4m+6n=r+s+2.$$
But it is easy to see, from the functional equation satisfied by $\zeta(s)$, viz.
$$(2\pi)^{-s}\Gamma(s)\zeta(s)\cos\tfrac{1}{2}\pi s=\tfrac{1}{2}\zeta(1-s) \quad\text{.....(39)},$$
that
$$F_{r,s}(0)=0 \quad\text{.....(40)}.$$

Hence $Q^3-R^2$ is a factor of the right-hand side in (38), that is to say
$$F_{r,s}(x)=(Q^3-R^2)\Sigma K_{m,n}Q^m R^n \quad\text{.....(41)},$$
where
$$4m+6n=r+s-10.$$

10. It is easy to deduce from (30) that
$$x\frac{d}{dx}\log(Q^3-R^2)=P \quad\text{.....(42)}.$$

But it is obvious that
$$P=x\frac{d}{dx}\log[x\{(1-x)(1-x^2)(1-x^3)\ldots\}^{24}] \quad\text{.....(43)};$$
and the coefficient of $x$ in $Q^3-R^2$ is 1728. Hence
$$Q^3-R^2=1728x\{(1-x)(1-x^2)(1-x^3)\ldots\}^{24} \quad\text{.....(44)}.$$

But it is known that
$$\{(1-x)(1-x^2)(1-x^3)(1-x^4)\ldots\}^3=1-3x+5x^3-7x^6+9x^{10}-\ldots \quad\text{.....(45)}.$$
Hence
$$Q^3-R^2=1728x(1-3x+5x^3-7x^6+\ldots)^8 \quad\text{.....(46)}.$$

The coefficient of $x^{\nu-1}$ in $1-3x+5x^3-\ldots$ is numerically less than $\sqrt{(8\nu)}$, and the coefficient of $x^\nu$ in $Q^3-R^2$ is therefore numerically less than that of $x^\nu$ in

$$1728x\{\sqrt{(8\nu)}(1+x+x^3+x^6+\ldots)\}^8.$$

But
$$x(1+x+x^3+x^6+\ldots)^8=\frac{1^3x}{1-x^2}+\frac{2^3x^2}{1-x^4}+\frac{3^3x^3}{1-x^6}+\ldots \quad\ldots\ldots\ldots\ldots(47),$$

and the coefficient of $x^\nu$ in the right-hand side is positive and less than

$$\nu^3\left(\frac{1}{1^3}+\frac{1}{3^3}+\frac{1}{5^3}+\ldots\right).$$

Hence the coefficient of $x^\nu$ in $Q^3-R^2$ is of the form

$$\nu^4 O(\nu^3)=O(\nu^7).$$

That is to say
$$Q^3-R^2=\Sigma\, O(\nu^7)x^\nu \quad\ldots\ldots\ldots\ldots(48).$$

Differentiating (48) and using (42) we obtain

$$P(Q^3-R^2)=\Sigma\, O(\nu^8)x^\nu \quad\ldots\ldots\ldots\ldots(49).$$

Differentiating this again with respect to $x$ we have

$$A(P^2-Q)(Q^3-R^2)+BQ(Q^3-R^2)=\Sigma\, O(\nu^9)x^\nu,$$

where $A$ and $B$ are constants. But

$$P^2-Q=-288\,\Phi_{1,2}(x)=-288\left\{\frac{1^2x}{(1-x)^2}+\frac{2^2x^2}{(1-x^2)^2}+\ldots\right\},$$

and the coefficient of $x^\nu$ in the right-hand side is a constant multiple of $\nu\sigma_1(\nu)$. Hence

$$(P^2-Q)(Q^3-R^2)=\Sigma\, O\,\nu\sigma_1(\nu)x^\nu\,\Sigma\, O(\nu^7)x^\nu$$
$$=\Sigma\, O(\nu^8)\{\sigma_1(1)+\sigma_1(2)+\ldots+\sigma_1(\nu)\}x^\nu=\Sigma\, O(\nu^{10})x^\nu,$$

and so
$$Q(Q^3-R^2)=\Sigma\, O(\nu^{10})x^\nu \quad\ldots\ldots\ldots\ldots(50).$$

Differentiating this again with respect to $x$ and using arguments similar to those used above, we deduce

$$R(Q^3-R^2)=\Sigma\, O(\nu^{12})x^\nu \quad\ldots\ldots\ldots\ldots(51).$$

Suppose now that $m$ and $n$ are any two positive integers including zero, and that $m+n$ is not zero. Then

$$Q^mR^n(Q^3-R^2)=Q(Q^3-R^2)Q^{m-1}R^n$$
$$=\Sigma\, O(\nu^{10})x^\nu\{\Sigma\, O(\nu^3)x^\nu\}^{m-1}\{\Sigma\, O(\nu^5)x^\nu\}^n$$
$$=\Sigma\, O(\nu^{10})x^\nu\,\Sigma\, O(\nu^{4m-5})x^\nu\,\Sigma\, O(\nu^{6n-1})x^\nu$$
$$=\Sigma\, O(\nu^{4m+6n+6})x^\nu,$$

if $m$ is not zero. Similarly we can show that

$$Q^mR^n(Q^3-R^2)=R(Q^3-R^2)Q^mR^n$$
$$=\Sigma\, O(\nu^{4m+6n-6})x^\nu,$$

if $n$ is not zero. Therefore in any case

$$(Q^3-R^2)Q^mR^n=\Sigma\, O(\nu^{4m-6n-6})x^\nu \quad\ldots\ldots\ldots\ldots(52).$$

11. Now let $r$ and $s$ be any two positive integers including zero. Then, when $r+s$ is equal to 2, 4, 6, 8 or 12, there are no values of $m$ and $n$ satisfying the relation

$$4m + 6n = r + s - 10$$

in (41); consequently in these cases

$$F_{r,s}(x) = 0 \qquad \text{.......(53).}$$

When $r+s=10$, $m$ and $n$ must both be zero, and this result does not apply; but it follows from (41) and (48) that

$$F_{r,s}(x) = \Sigma\, O(\nu^7)\, x^\nu \qquad \text{.......(54).}$$

And when $r+s \geqslant 14$ it follows from (52) that

$$F_{r,s}(x) = \Sigma\, O(\nu^{r+s-4})\, x^\nu \qquad \text{.......(55).}$$

Equating the coefficients of $x^\nu$ in both sides in (53), (54) and (55) we obtain

$$\Sigma_{r,s}(n) = \frac{\Gamma(r+1)\,\Gamma(s+1)}{\Gamma(r+s+2)} \frac{\zeta(r+1)\,\zeta(s+1)}{\zeta(r+s+2)} \sigma_{r+s+1}(n) + \frac{\zeta(1-r)+\zeta(1-s)}{r+s}\, n\,\sigma_{r+s-1}(n) + E_{r,s}(n) \qquad \text{......(56),}$$

where

$$E_{r,s}(n) = 0, \quad r+s = 2,\ 4,\ 6,\ 8,\ 12;$$

$$E_{r,s}(n) = O(n^7), \quad r+s = 10;$$

$$E_{r,s}(n) = O(n^{r+s-4}), \quad r+s \geqslant 14.$$

Since $\sigma_{r+s+1}(n)$ is of order $n^{r+s+1}$ it follows that in all cases

$$\Sigma_{r,s}(n) \sim \frac{\Gamma(r+1)\,\Gamma(s+1)}{\Gamma(r+s+2)} \frac{\zeta(r+1)\,\zeta(s+1)}{\zeta(r+s+2)} \sigma_{r+s+1}(n) \qquad \text{.......(57).}$$

The following table gives the values of $\Sigma_{r,s}(n)$ when $r+s = 2,\ 4,\ 6,\ 8,\ 12$.

TABLE IV.

1. $\Sigma_{1,1}(n) = \dfrac{5\sigma_3(n) - 6n\sigma_1(n)}{12}.$

2. $\Sigma_{1,3}(n) = \dfrac{7\sigma_5(n) - 10n\sigma_3(n)}{80}.$

3. $\Sigma_{3,3}(n) = \dfrac{\sigma_7(n)}{120}.$

4. $\Sigma_{1,5}(n) = \dfrac{10\sigma_7(n) - 21n\sigma_5(n)}{252}.$

5. $\Sigma_{3,5}(n) = \dfrac{11\sigma_9(n)}{5040}.$

6. $\Sigma_{1,7}(n) = \dfrac{11\sigma_9(n) - 30n\sigma_7(n)}{480}.$

7. $\Sigma_{5,7}(n) = \dfrac{\sigma_{13}(n)}{10080}.$

8. $\Sigma_{3,9}(n) = \dfrac{\sigma_{13}(n)}{2640}.$

9. $\Sigma_{1,11}(n) = \dfrac{691\sigma_{13}(n) - 2730n\sigma_{11}(n)}{65520}.$

12. In this connection it may be interesting to note that

$$\sigma_1(1)\sigma_3(n)+\sigma_1(3)\sigma_3(n-1)+\sigma_1(5)\sigma_3(n-2)+\ldots+\sigma_1(2n+1)\sigma_3(0)=\tfrac{1}{240}\sigma_5(2n+1) \quad \ldots\ldots\ldots(58).$$

This formula may be deduced from the identity

$$\frac{1^5x}{1-x}+\frac{3^5x^2}{1-x^3}+\frac{5^5x^3}{1-x^5}+\ldots=Q\left(\frac{x}{1-x}+\frac{3x^2}{1-x^3}+\frac{5x^3}{1-x^5}+\ldots\right) \quad \ldots\ldots\ldots\ldots(59),$$

which can be proved by means of the theory of elliptic functions or by elementary methods.

13. More precise results concerning the order of $E_{r,s}(n)$ can be deduced from the theory of elliptic functions. Let

$$x=q^2.$$

Then we have

$$\left.\begin{aligned} Q&=\phi^8(q)\{1-(kk')^2\}\\ R&=\phi^{12}(q)(k'^2-k^2)\{1+\tfrac{1}{2}(kk')^2\}\\ &=\phi^{12}(q)\{1+\tfrac{1}{2}(kk')^2\}\sqrt{\{1-(2kk')^2\}}\end{aligned}\right\} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots(60),$$

where $\qquad \phi(q)=1+2q+2q^4+2q^9+\ldots.$

But if $\qquad f(q)=q^{\frac{1}{24}}(1-q)(1-q^2)(1-q^3)\ldots,$

then we know that

$$\left.\begin{aligned} 2^{\frac{1}{6}}f(q)&=k^{\frac{1}{12}}k'^{\frac{1}{3}}\phi(q)\\ 2^{\frac{1}{6}}f(-q)&=(kk')^{\frac{1}{12}}\phi(q)\\ 2^{\frac{1}{3}}f(q^2)&=(kk')^{\frac{1}{6}}\phi(q)\\ 2^{\frac{2}{3}}f(q^4)&=k^{\frac{1}{3}}k'^{\frac{1}{12}}\phi(q)\end{aligned}\right\} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots(61).$$

It follows from (41), (60) and (61) that, if $r+s$ is of the form $4m+2$, but not equal to 2 or to 6, then

$$F_{r,s}(q^2)=\frac{f^{4(r+s-4)}(-q)}{f^{2(r+s-10)}(q^2)}\sum_1^{\frac{1}{4}(r+s-6)}K_n\frac{f^{24n}(q^2)}{f^{24n}(-q)} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots(62),$$

and if $r+s$ is of the form $4m$, but not equal to 4, 8 or 12, then

$$F_{r,s}(q^2)=\frac{f^{4(r+s-6)}(-q)}{f^{2(r+s-10)}(q^2)}\{f^8(q)-16f^8(q^4)\}\sum_1^{\frac{1}{4}(r+s-8)}K_n\frac{f^{24n}(q^2)}{f^{24n}(-q)} \quad \ldots\ldots\ldots\ldots(63),$$

where $K_n$ depends on $r$ and $s$ only. Hence it is easy to see that in all cases $F_{r,s}(q^2)$ can be expressed as

$$\Sigma K_{a,b,c,d,e,h,k}\{f^3(-q)\}^a\left\{\frac{f^5(-q)}{f^2(q^2)}\right\}^b\left\{\frac{f^5(q^2)}{f^2(-q)}\right\}^c\left\{\frac{f^5(q)}{f^2(q^2)}f^3(q)\right\}^d\left\{\frac{f^5(q^4)}{f^2(q^2)}f^3(q^4)\right\}^e f^h(-q)f^k(q^2) \quad \ldots\ldots\ldots(64),$$

where $a$, $b$, $c$, $d$, $e$, $h$, $k$ are zero or positive integers such that

$$a+b+c+2(d+e)=[\tfrac{2}{3}(r+s+2)],$$

$$h+k=2(r+s+2)-3[\tfrac{2}{3}(r+s+2)],$$

and $[x]$ denotes as usual the greatest integer in $x$. But

$$\left.\begin{aligned} f(q) &= q^{\frac{1^2}{24}} - q^{\frac{5^2}{24}} - q^{\frac{7^2}{24}} + q^{\frac{11^2}{24}} + \dots \\ f^3(q) &= q^{\frac{1^2}{8}} - 3q^{\frac{3^2}{8}} + 5q^{\frac{5^2}{8}} - 7q^{\frac{7^2}{8}} + \dots \\ \frac{f^5(q)}{f^2(q^2)} &= q^{\frac{1^2}{24}} - 5q^{\frac{5^2}{24}} + 7q^{\frac{7^2}{24}} - 11q^{\frac{11^2}{24}} + \dots \\ \frac{f^5(q^2)}{f^2(-q)} &= q^{\frac{1^2}{3}} - 2q^{\frac{2^2}{3}} + 4q^{\frac{4^2}{3}} - 5q^{\frac{5^2}{3}} + \dots \end{aligned}\right\} \quad \dots\dots\dots\dots(65),$$

where 1, 2, 4, 5, ... are the natural numbers without the multiples of 3, and 1, 5, 7, 11, ... are the natural odd numbers without the multiples of 3.

Hence it is easy to see that

$$n^{-\frac{1}{2}(a+b+c)-d-e}\, E_{r,s}(n)$$

is not of higher order than the coefficient of $q^{2n}$ in

$$\phi^a(q^{\frac{1}{8}})\,\phi^b(q^{\frac{1}{24}})\,\phi^c(q^{\frac{1}{3}})\,\{\phi(q^{\frac{1}{24}})\,\phi(q^{\frac{1}{8}})\}^d\,\{\phi(q^{\frac{2}{3}})\,\phi(q^{\frac{1}{2}})\}^e\,\phi^h(q^{\frac{1}{24}})\,\phi^k(q^{\frac{1}{12}}),$$

or the coefficient of $q^{48n}$ in

$$\phi^{a+d}(q^3)\,\phi^{b+d+h}(q)\,\phi^c(q^8)\,\phi^e(q^{16})\,\phi^e(q^{12})\,\phi^k(q^2).$$

But the coefficient of $q^\nu$ in $\phi^2(q^2)$ cannot exceed that of $q^\nu$ in $\phi^2(q)$, since

$$\phi^2(q) + \phi^2(-q) = 2\phi^2(q^2) \quad \dots\dots\dots\dots(66);$$

and it is evident that the coefficient of $q^\nu$ in $\phi(q^{4\lambda})$ cannot exceed that of $q^\nu$ in $\phi(q^\lambda)$. Hence it follows that

$$n^{-\frac{1}{2}[\frac{2}{3}(r+s+2)]}\, E_{r,s}(n)$$

is not of higher order than the coefficient of $q^{48n}$ in

$$\phi^A(q)\,\phi^B(q^3)\,\phi^C(q^2),$$

where $A$, $B$, $C$ are zero or positive integers such that

$$A + B + C = 2(r+s+2) - 2\,[\tfrac{2}{3}(r+s+2)],$$

and $C$ is 0 or 1.

Now, if $r + s \geqslant 14$, we have $\quad A + B + C \geqslant 12,$

and so $\quad A + B \geqslant 11.$

Therefore one at least of $A$ and $B$ is greater than 5. But

$$\phi^6(q) = \sum_0^\infty O(\nu^2)\, q^\nu\,* \quad \dots\dots\dots\dots(67).$$

Hence it is easily deduced that

$$\phi^A(q)\,\phi^B(q^3)\,\phi^C(q^2) = \Sigma\, O\{\nu^{\frac{1}{2}(A+B+C)-1}\}\, q^\nu \quad \dots\dots\dots\dots(68).$$

It follows that

$$E_{r,s}(n) = O\{n^{r+s-\frac{1}{2}[\frac{2}{3}(r+s-1)]}\} \quad \dots\dots\dots\dots(69),$$

if $r + s \geqslant 14$. We have already shown in § 11 that, if $r + s = 10$, then

$$E_{r,s}(n) = O(n^7) \quad \dots\dots\dots\dots(70).$$

* See §§ 24—25.

This agrees with (69). Thus we see that in all cases

$$E_{r,s}(n) = O\{n^{\frac{2}{3}(r+s+1)}\} \quad \text{.....(71)};$$

and that, if $r+s$ is of the form $6m$, then

$$E_{r,s}(n) = O\{n^{\frac{2}{3}(r+s+\frac{3}{4})}\} \quad \text{.....(72)},$$

and if of the form $6m+4$, then

$$E_{r,s}(n) = O\{n^{\frac{2}{3}(r+s+\frac{1}{2})}\} \quad \text{.....(73)}.$$

14. I shall now prove that the order of $E_{r,s}(n)$ is not less than that of $n^{\frac{1}{2}(r+s)}$. In order to prove this result I shall follow the method used by Messrs Hardy and Littlewood in their paper 'Some problems of Diophantine approximation' (II)*.

Let

$$q = e^{\pi i\tau}, \quad q' = e^{\pi i T},$$

where

$$T = \frac{c+d\tau}{a+b\tau},$$

and

$$ad - bc = 1.$$

Also let

$$V = \frac{v}{a+b\tau}.$$

Then we have

$$\omega\sqrt{v}e^{\pi ibvV}\vartheta_1(v,\tau) = \sqrt{V}\vartheta_1(V,T) \quad \text{.....(74)},$$

where $\omega$ is an eighth root of unity and

$$\vartheta_1(v,\tau) = 2\sin\pi v \,.\, q^{\frac{1}{4}}\prod_1^\infty(1-q^{2n})(1-2q^{2n}\cos 2\pi v + q^{4n}) \quad \text{.....(75)}.$$

From (75) we have

$$\log\vartheta_1(v,\tau) = \log(2\sin\pi v) + \tfrac{1}{4}\log q - \sum_1^\infty \frac{q^{2n}(1+2\cos 2n\pi v)}{n(1-q^{2n})} \quad \text{.....(76)}.$$

It follows from (74) and (76) that

$$\log\sin\pi v + \tfrac{1}{2}\log v + \tfrac{1}{4}\log q + \log\omega - \sum_1^\infty \frac{q^{2n}(1+2\cos 2n\pi v)}{n(1-q^{2n})}$$

$$= \log\sin\pi V + \tfrac{1}{2}\log V + \tfrac{1}{4}\log q' - \pi ibvV - \sum_1^\infty \frac{q'^{2n}(1+2\cos 2n\pi V)}{n(1-q'^{2n})} \quad \text{.....(77)}.$$

Equating the coefficients of $v^{s+1}$ on the two sides of (77), we obtain

$$(a+b\tau)^{s+1}\left\{\tfrac{1}{2}\zeta(-s) + \frac{1^s q^2}{1-q^2} + \frac{2^s q^4}{1-q^4} + \frac{3^s q^6}{1-q^6} + \dots\right\}$$

$$= \tfrac{1}{2}\zeta(-s) + \frac{1^s q'^2}{1-q'^2} + \frac{2^s q'^4}{1-q'^4} + \frac{3^s q'^6}{1-q'^6} + \dots \quad \text{.....(78)},$$

provided that $s$ is an odd integer greater than 1. If, in particular, we put $s=3$ and $s=5$ in (78) we obtain

$$(a+b\tau)^4\left\{1 + 240\left(\frac{1^3 q^2}{1-q^2} + \frac{2^3 q^4}{1-q^4} + \frac{3^3 q^6}{1-q^6} + \dots\right)\right\}$$

$$= \left\{1 + 240\left(\frac{1^3 q'^2}{1-q'^2} + \frac{2^3 q'^4}{1-q'^4} + \frac{3^3 q'^6}{1-q'^6} + \dots\right)\right\} \quad \text{.....(79)},$$

and

$$(a+b\tau)^6\left\{1 - 504\left(\frac{1^5 q^2}{1-q^2} + \frac{2^5 q^4}{1-q^4} + \frac{3^5 q^6}{1-q^6} + \dots\right)\right\}$$

$$= \left\{1 - 504\left(\frac{1^5 q'^2}{1-q'^2} + \frac{2^5 q'^4}{1-q'^4} + \frac{3^5 q'^6}{1-q'^6} + \dots\right)\right\} \quad \text{.....(80)}$$

* *Acta Mathematica*, Vol. XXXVII. pp. 193—238.

It follows from (38), (79) and (80) that

$$(a+b\tau)^{r+s+2} F_{r,s}(q^2) = F_{r,s}(q'^2) \quad \text{.............................(81).}$$

It can easily be seen from (56) and (37) that

$$F_{r,s}(x) = \sum_1^\infty E_{r,s}(n)\, x^n \quad \text{.....................................(82).}$$

Hence

$$(a+b\tau)^{r+s+2} \sum_1^\infty E_{r,s}(n)\, q^{2n} = \sum_1^\infty E_{r,s}(n)\, q'^{2n} \quad \text{........................(83).}$$

It is important to observe that

$$E_{r,s}(1) = \frac{\zeta(-r)+\zeta(-s)}{2} - \frac{\zeta(1-r)+\zeta(1-s)}{r+s} - \frac{\Gamma(r+1)\,\Gamma(s+1)}{\Gamma(r+s+2)}\frac{\zeta(r+1)\,\zeta(s+1)}{\zeta(r+s+2)}$$

$$\neq 0 \quad \text{..............................................................................(84),}$$

if $r+s$ is not equal to 2, 4, 6, 8 or 12. This is easily proved by the help of the equation (39).

15. Now let

$$\tau = u+iy, \quad t = e^{-\pi y} \quad (u>0,\ y>0,\ 0<t<1),$$

so that

$$q = e^{\pi iu - \pi y} = te^{\pi iu};$$

and let us suppose that $p_n/q_n$ is a convergent to

$$u = \frac{1}{a_1+}\,\frac{1}{a_2+}\,\frac{1}{a_3}+\ldots,$$

so that

$$\eta_n = p_{n-1}q_n - p_nq_{n-1} = \pm 1.$$

Further, let us suppose that

$$a = p_n, \qquad b = -q_n,$$
$$c = \eta_n p_{n-1}, \quad d = -\eta_n q_{n-1},$$

so that

$$ad-bc = \eta_n^2 = 1.$$

Furthermore, let

$$y = 1/(q_n q'_{n+1}),$$

where

$$q'_{n+1} = a'_{n+1}q_n + q_{n-1},$$

and $a'_{n+1}$ is the complete quotient corresponding to $a_{n+1}$.

Then we have

$$a+b\tau = p_n - q_n u - iq_n y = \frac{\pm 1 - i}{q'_{n+1}} = \frac{\sqrt{2}}{q'_{n+1}} \quad \text{....................(85),}$$

and

$$q' = e^{-\pi\lambda},$$

where

$$\lambda = \mathbf{I}(T) = \mathbf{I}\left(\frac{c+d\tau}{a+b\tau}\right) = \mathbf{I}\left\{\frac{d}{b} - \frac{1}{b(a+b\tau)}\right\}$$

$$= \frac{y}{(1/q'_{n+1})^2 + q_n^2 y^2} = \frac{q'_{n+1}}{2q_n} \quad \text{................................(86),}$$

and $\mathbf{I}(T)$ is the imaginary part of $T$. It follows from (83), (85) and (86) that

$$\left|\sum_1^\infty E_{r,s}(n)\, q^{2n}\right| = \left(\frac{q'_{n+1}}{\sqrt{2}}\right)^{r+s+2} \left|\sum_1^\infty E_{r,s}(n)\, q'^{2n}\right|$$

$$\geq \left(\frac{q'_{n+1}}{\sqrt{2}}\right)^{r+s+2} \{E_{r,s}(1)\, e^{-2\pi\lambda} - E_{r,s}(2)\, e^{-4\pi\lambda} - E_{r,s}(3)\, e^{-6\pi\lambda} - \ldots\} \text{...(87).}$$

We can choose a number $\lambda_0$, depending only on $r$ and $s$, such that

$$|E_{r,s}(1)|\, e^{-2\pi\lambda} > 2\{|E_{r,s}(2)|\, e^{-4\pi\lambda} + |E_{r,s}(3)|\, e^{-6\pi\lambda} + \ldots\}$$

for $\lambda \geqslant \lambda_0$. Let us suppose $\lambda_0 > 10$. Let us also suppose that the continued fraction for $u$ satisfies the condition

$$4\lambda_0 q_n > q'_{n+1} > 2\lambda_0 q_n \qquad \ldots\ldots(88)$$

for an infinity of values of $n$. Then

$$\left|\sum_1^\infty E_{r,s}(n)\, q^{2n}\right| \geqslant \tfrac{1}{2}\, |E_{r,s}(1)| \left(\frac{q'_{n+1}}{\sqrt{2}}\right)^{r+s+2} e^{-4\pi\lambda_0} > K\,(q'_{n+1})^{r+s+2} \qquad \ldots\ldots(89),$$

where $K$ depends on $r$ and $s$ only. Also

$$q_n q'_{n+1} = 1/y,$$

$$q'_{n+1} > \frac{1}{\sqrt{y}} = \sqrt{\left\{\frac{\pi}{\log(1/t)}\right\}} > \frac{K}{\sqrt{(1-t)}}.$$

It follows that, if $u$ is an irrational number such that the condition (88) is satisfied for an infinity of values of $n$, then

$$\left|\sum_1^\infty E_{r,s}(n)\, q^{2n}\right| > K\,(1-t)^{-\frac{1}{2}(r+s+2)} \qquad \ldots\ldots(90)$$

for an infinity of values of $t$ tending to unity.

But if we had $\qquad E_{r,s}(n) = o\,\{n^{\frac{1}{2}(r+s)}\}$,

then we should have $\qquad \left|\sum_1^\infty E_{r,s}(n)\, q^{2n}\right| = o\,\{(1-t)^{-\frac{1}{2}(r+s+2)}\}$,

which contradicts (90). It follows that the error term in $\Sigma_{r,s}(n)$ is not of the form

$$o\,\{n^{\frac{1}{2}(r+s)}\} \qquad \ldots\ldots(91).$$

*The arithmetical function $\tau(n)$.*

16. We have seen that $\qquad E_{r,s}(n) = 0,$

if $r+s$ is equal to 2, 4, 6, 8 or 12. In these cases $\Sigma_{r,s}(n)$ has a finite expression in terms of $\sigma_{r+s+1}(n)$ and $\sigma_{r+s-1}(n)$. In other cases $\Sigma_{r,s}(n)$ involves other arithmetical functions as well. The simplest of these is the function $\tau(n)$ defined by

$$\sum_1^\infty \tau(n)\, x^n = x\,\{(1-x)(1-x^2)(1-x^3)\ldots\}^{24} \qquad \ldots\ldots(92).$$

These cases arise when $r+s$ has one of the values 10, 14, 16, 18, 20 or 24.

Suppose that $r+s$ has one of these values. Then

$$\frac{1728 \sum_1^\infty E_{r,s}(n)\, x^n}{(Q^3 - R^2)\, E_{r,s}(1)}$$

is, by (41) and (82), equal to the corresponding one of the functions

$$1,\ Q,\ R,\ Q^2,\ QR,\ Q^2R.$$

In other words

$$\sum_1^\infty E_{r,s}(n)\, x^n = E_{r,s}(1) \sum_1^\infty \tau(n)\, x^n \left\{1 + \frac{2}{\zeta(11-r-s)} \sum_1^\infty n^{r+s-11} \frac{x^n}{1-x^n}\right\} \qquad \ldots\ldots(93).$$

We thus deduce the formulae

$$E_{r,s}(n) = E_{r,s}(1)\,\tau(n) \qquad\qquad (94),$$

if $r+s=10$; and

$$\sigma_{r+s-11}(0)\,E_{r,s}(n)$$
$$= E_{r,s}(1)\,\{\sigma_{r+s-11}(0)\,\tau(n) + \sigma_{r+s-11}(1)\,\tau(n-1) + \ldots + \sigma_{r+s-11}(n-1)\,\tau(1)\} \qquad (95),$$

if $r+s$ is equal to 14, 16, 18, 20 or 24. It follows from (94) and (95) that, if $r+s=r'+s'$, then

$$E_{r,s}(n)\,E_{r',s'}(1) = E_{r,s}(1)\,E_{r',s'}(n) \qquad\qquad (96),$$

and in general

$$E_{r,s}(m)\,E_{r',s'}(n) = E_{r,s}(n)\,E_{r',s'}(m) \qquad\qquad (97),$$

when $r+s$ has one of the values in question. The different cases in which $r+s$ has the same value are therefore not fundamentally distinct.

17. The values of $\tau(n)$ may be calculated as follows: differentiating (92) logarithmically with respect to $x$, we obtain

$$\sum_1^\infty n\,\tau(n)\,x^n = P\sum_1^\infty \tau(n)\,x^n \qquad\qquad (98).$$

Equating the coefficients of $x^n$ in both sides in (98), we have

$$\tau(n) = \frac{24}{1-n}\{\sigma_1(1)\,\tau(n-1) + \sigma_1(2)\,\tau(n-2) + \ldots + \sigma_1(n-1)\,\tau(1)\} \qquad (99).$$

If, instead of starting with (92), we start with

$$\sum_1^\infty \tau(n)\,x^n = x(1-3x+5x^3-7x^6+\ldots)^8,$$

we can show that

$$(n-1)\,\tau(n) - 3(n-10)\,\tau(n-1) + 5(n-28)\,\tau(n-3) - 7(n-55)\,\tau(n-6)$$
$$+ \ldots \text{ to } [\tfrac{1}{2}\{1+\sqrt{(8n-7)}\}] \text{ terms} = 0 \qquad\qquad (100),$$

where the $r$th term of the sequence 0, 1, 3, 6, ... is $\frac{1}{2}r(r-1)$, and the $r$th term of the sequence 1, 10, 28, 55, ... is $1+\frac{9}{2}r(r-1)$. We thus obtain the values of $\tau(n)$ in the following table.

TABLE V.

| $n$ | $\tau(n)$ | $n$ | $\tau(n)$ |
|---|---|---|---|
| 1 | + 1 | 16 | + 987136 |
| 2 | − 24 | 17 | − 6905934 |
| 3 | + 252 | 18 | + 2727432 |
| 4 | − 1472 | 19 | + 10661420 |
| 5 | + 4830 | 20 | − 7109760 |
| 6 | − 6048 | 21 | − 4219488 |
| 7 | − 16744 | 22 | − 12830688 |
| 8 | + 84480 | 23 | + 18643272 |
| 9 | − 113643 | 24 | + 21288960 |
| 10 | − 115920 | 25 | − 25499225 |
| 11 | + 534612 | 26 | + 13865712 |
| 12 | − 370944 | 27 | − 73279080 |
| 13 | − 577738 | 28 | + 24647168 |
| 14 | + 401856 | 29 | + 128406630 |
| 15 | + 1217160 | 30 | − 29211840 |

18. Let us consider more particularly the case in which $r+s=10$. The order of $E_{r,s}(n)$ is then the same as that of $\tau(n)$. The determination of this order is a problem interesting in itself. We have proved that $E_{r,s}(n)$, and therefore $\tau(n)$, is of the form $O(n^7)$ and not of the form $o(n^5)$. There is reason for supposing that $\tau(n)$ is of the form $O(n^{\frac{11}{2}+\epsilon})$ and not of the form $o(n^{\frac{11}{2}})$. For it appears that

$$\sum_1^\infty \frac{\tau(n)}{n^t} = \prod_p \frac{1}{1-\tau(p)\,p^{-t}+p^{11-2t}} \quad \text{.........................(101).}$$

This assertion is equivalent to the assertion that, if

$$n = p_1^{a_1} p_2^{a_2} p_3^{a_3} \dots p_r^{a_r},$$

where $p_1, p_2, \dots p_r$ are the prime divisors of $n$, then

$$n^{-\frac{11}{2}}\,\tau(n) = \frac{\sin(1+a_1)\,\theta_{p_1}}{\sin\theta_{p_1}} \frac{\sin(1+a_2)\,\theta_{p_2}}{\sin\theta_{p_2}} \dots \frac{\sin(1+a_r)\,\theta_{p_r}}{\sin\theta_{p_r}} \quad \text{...........(102),}$$

where $$\cos\theta_p = \tfrac{1}{2}p^{-\frac{11}{2}}\,\tau(p).$$

It would follow that, if $n$ and $n'$ are prime to each other, we must have

$$\tau(nn') = \tau(n)\,\tau(n') \quad \text{..................................(103).}$$

Let us suppose that (102) is true, and also that (as appears to be highly probable)

$$\{2\tau(p)\}^2 \leqslant p^{11} \quad \text{....................................(104),}$$

so that $\theta_p$ is real. Then it follows from (102) that

$$n^{-\frac{11}{2}}\,\tau(n) \leqslant (1+a_1)(1+a_2)\dots(1+a_r),$$

that is to say $$|\tau(n)| \leqslant n^{\frac{11}{2}}\,d(n) \quad \text{..................................(105),}$$

where $d(n)$ denotes the number of divisors of $n$.

Now let us suppose that $n=p^a$, so that

$$n^{-\frac{11}{2}}\,\tau(n) = \frac{\sin(1+a)\,\theta_p}{\sin\theta_p}.$$

Then we can choose $a$ as large as we please and such that

$$\left|\frac{\sin(1+a)\,\theta_p}{\sin\theta_p}\right| \geqslant 1.$$

Hence $$|\tau(n)| \geqslant n^{\frac{11}{2}} \quad \text{.........................................(106)}$$

for an infinity of values of $n$.

19. It should be observed that precisely similar questions arise with regard to the arithmetical function $\psi(n)$ defined by

$$\sum_0^\infty \psi(n)\,x^n = f^{a_1}(x^{c_1})\,f^{a_2}(x^{c_2})\dots f^{a_r}(x^{c_r}) \quad \text{......................(107),}$$

where $$f(x) = x^{\frac{1}{24}}(1-x)(1-x^2)(1-x^3)\dots,$$

the $a$'s and $c$'s are integers, the latter being positive,

$$\tfrac{1}{24}(a_1c_1 + a_2c_2 + \ldots + a_rc_r)$$

is equal to 0 or 1, and

$$l\left(\frac{a_1}{c_1} + \frac{a_2}{c_2} + \ldots + \frac{a_r}{c_r}\right),$$

where $l$ is the least common multiple of $c_1, c_2, \ldots c_r$, is equal to 0 or to a divisor of 24.

The arithmetical functions $\chi(n)$, $P(n)$, $\chi_4(n)$, $\Omega(n)$ and $\Theta(n)$, studied by Dr Glaisher in the *Quarterly Journal*, Vols. XXXVI.—XXXVIII., are of this type. Thus

$$\sum_1^\infty \chi(n)\,x^n = f^6(x^4),$$

$$\sum_1^\infty P(n)\,x^n = f^4(x^2)f^4(x^4),$$

$$\sum_1^\infty \chi_4(n)\,x^n = f^4(x)f^2(x^2)f^4(x^4),$$

$$\sum_1^\infty \Omega(n)\,x^n = f^{12}(x^2),$$

$$\sum_1^\infty \Theta(n)\,x^n = f^8(x)f^8(x^2).$$

20. The results (101) and (104) may be written as

$$\sum_1^\infty \frac{E_{r,s}(n)}{n^t} = E_{r,s}(1)\prod_p \frac{1}{1-2c_p\,p^{-t}+p^{r+s+1-2t}} \quad \ldots\ldots\ldots\ldots(108),$$

where

$$c_p^2 \leqslant p^{r+s+1},$$

and

$$2c_p E_{r,s}(1) = E_{r,s}(p).$$

It seems probable that the result (108) is true not only for $r+s=10$ but also when $r+s$ is equal to 14, 16, 18, 20 or 24, and that

$$\left|\frac{E_{r,s}(n)}{E_{r,s}(1)}\right| \leqslant n^{\frac{1}{2}(r+s+1)}\,d(n) \quad \ldots\ldots\ldots\ldots(109)$$

for all values of $n$, and

$$\left|\frac{E_{r,s}(n)}{E_{r,s}(1)}\right| \geqslant n^{\frac{1}{2}(r+s+1)} \quad \ldots\ldots\ldots\ldots(110)$$

for an infinity of values of $n$. If this be so then

$$E_{r,s}(n) = O\{n^{\frac{1}{2}(r+s+1+\epsilon)}\}, \quad E_{r,s}(n) \neq o\{n^{\frac{1}{2}(r+s+1)}\} \quad \ldots\ldots\ldots\ldots(111).$$

And it seems very likely that these equations hold generally, whenever $r$ and $s$ are positive odd integers.

21. It is of some interest to see what confirmation of these conjectures can be found from a study of the coefficients in the expansion of

$$x\{(1-x^{24/\alpha})(1-x^{48/\alpha})(1-x^{72/\alpha})\ldots\}^\alpha = \sum_1^\infty \psi_\alpha(n)\,x^n,$$

where $\alpha$ is a divisor of 24. When $\alpha = 1$ and $\alpha = 3$ we know the actual value of $\psi_\alpha(n)$. For we have

$$\sum_1^\infty \psi_1(n) x^n = x^{1^2} - x^{5^2} - x^{7^2} + x^{11^2} + x^{13^2} - x^{17^2} + \dots \quad \text{.............(112)},$$

where 1, 5, 7, 11, ... are the natural odd numbers without the multiples of 3; and

$$\sum_1^\infty \psi_3(n) x^n = x^{1^2} - 3x^{3^2} + 5x^{5^2} - 7x^{7^2} + \dots \quad \text{.......................(113)}.$$

The corresponding Dirichlet's series are

$$\sum_1^\infty \frac{\psi_1(n)}{n^s} = \frac{1}{(1+5^{-2s})(1+7^{-2s})(1-11^{-2s})(1-13^{-2s})\dots} \quad \text{.............(114)},$$

where 5, 7, 11, 13, ... are the primes greater than 3, those of the form $12n \pm 5$ having the plus sign and those of the form $12n \pm 1$ the minus sign; and

$$\sum_1^\infty \frac{\psi_3(n)}{n^s} = \frac{1}{(1+3^{1-2s})(1-5^{1-2s})(1+7^{1-2s})(1+11^{1-2s})\dots} \quad \text{.............(115)},$$

where 3, 5, 7, 11, ... are the odd primes, those of the form $4n-1$ having the plus sign and those of the form $4n+1$ the minus sign.

It is easy to see that

$$|\psi_1(n)| \leqslant 1, \quad |\psi_3(n)| \leqslant \sqrt{n} \quad \text{...............................(116)}$$

for all values of $n$, and

$$|\psi_1(n)| = 1, \quad |\psi_3(n)| = \sqrt{n} \quad \text{...............................(117)}$$

for an infinity of values of $n$.

The next simplest case is that in which $\alpha = 2$. In this case it appears that

$$\sum_1^\infty \frac{\psi_2(n)}{n^s} = \Pi_1 \Pi_2 \quad \text{.....................................(118)},$$

where
$$\Pi_1 = \frac{1}{(1+5^{-2s})(1-7^{-2s})(1-11^{-2s})(1+17^{-2s})\dots},$$

5, 7, 11, ... being the primes of the forms $12n-1$ and $12n \pm 5$, those of the form $12n+5$ having the plus sign and the rest the minus sign; and

$$\Pi_2 = \frac{1}{(1+13^{-s})^2(1-37^{-s})^2(1-61^{-s})^2(1+73^{-s})^2\dots},$$

13, 37, 61, ... being the primes of the form $12n+1$, those of the form $m^2+(6n-3)^2$ having the plus sign and those of the form $m^2+(6n)^2$ the minus sign.

This is equivalent to the assertion that if

$$n = (5^{a_5} . 7^{a_7} . 11^{a_{11}} . 17^{a_{17}} \dots)^2 13^{a_{13}} . 37^{a_{37}} . 61^{a_{61}} . 73^{a_{73}} \dots,$$

where $a_p$ is zero or a positive integer, then

$$\psi_2(n) = (-1)^{a_5 + a_{13} + a_{17} + a_{29} + a_{41} + \dots} (1+a_{13})(1+a_{37})(1+a_{61}) \dots \quad \text{........(119)},$$

where 5, 13, 17, 29, ... are the primes of the form $4n+1$, excluding those of the form $m^2+(6n)^2$; and that otherwise

$$\psi_2(n) = 0 \qquad \text{.....(120).}$$

It follows that

$$|\psi_2(n)| \leqslant d(n) \qquad \text{.....(121)}$$

for all values of $n$, and

$$|\psi_2(n)| \geqslant 1 \qquad \text{.....(122)}$$

for an infinity of values of $n$. These results are easily proved to be actually true.

22. I have investigated also the cases in which $\alpha$ has one of the values 4, 6, 8 or 12. Thus for example, when $\alpha = 6$, I find

$$\sum_1^\infty \frac{\psi_6(n)}{n^s} = \Pi_1 \Pi_2{}^* \qquad \text{.....(123),}$$

where

$$\Pi_1 = \frac{1}{(1-3^{2-2s})(1-7^{2-2s})(1-11^{2-2s})\dots},$$

3, 7, 11, ... being the primes of the form $4n-1$; and

$$\Pi_2 = \frac{1}{(1-2c_5 \,.\, 5^{-s}+5^{2-2s})(1-2c_{13} \,.\, 13^{-s}+13^{2-2s})\dots},$$

5, 13, 17, ... being the primes of the form $4n+1$, and $c_p = u^2-(2v)^2$, where $u$ and $v$ are the unique pair of positive integers for which $p = u^2+(2v)^2$. This is equivalent to the assertion that if

$$n = (3^{a_3} \,.\, 7^{a_7} \,.\, 11^{a_{11}} \dots)^2 \,.\, 5^{a_5} \,.\, 13^{a_{13}} \,.\, 17^{a_{17}} \dots,$$

then

$$\frac{\psi_6(n)}{n} = \frac{\sin(1+a_5)\theta_5}{\sin\theta_5} \,.\, \frac{\sin(1+a_{13})\theta_{13}}{\sin\theta_{13}} \,.\, \frac{\sin(1+a_{17})\theta_{17}}{\sin\theta_{17}} \dots \qquad \text{.....(124),}$$

where

$$\tan \tfrac{1}{2}\theta_p = \frac{u}{2v} \qquad (0 < \theta_p < \pi),$$

and that otherwise $\psi_6(n) = 0$. From these results it would follow that

$$|\psi_6(n)| \leqslant n\,d(n) \qquad \text{.....(125)}$$

for all values of $n$, and

$$|\psi_6(n)| \geqslant n \qquad \text{.....(126)}$$

for an infinity of values of $n$. What can actually be proved to be true is that

$$|\psi_6(n)| < 2n\,d(n)$$

for all values of $n$, and

$$|\psi_6(n)| \geqslant n$$

for an infinity of values of $n$.

23. In the case in which $\alpha = 4$ I find that, if

$$n = (5^{a_5} \,.\, 11^{a_{11}} \,.\, 17^{a_{17}} \dots)^2 \,.\, 7^{a_7} \,.\, 13^{a_{13}} \,.\, 19^{a_{19}} \dots,$$

* $\psi_6(n)$ is Dr Glaisher's $\lambda(n)$.

where 5, 11, 17, ... are the primes of the form $6m-1$ and 7, 13, 19, ... are those of the form $6m+1$, then

$$\frac{\psi_4(n)}{\sqrt{n}} = (-1)^{a_5+a_{11}+a_{17}+\dots} \frac{\sin(1+a_7)\theta_7}{\sin\theta_7} \cdot \frac{\sin(1+a_{13})\theta_{13}}{\sin\theta_{13}} \dots \quad \text{.........(127)},$$

where

$$\tan\theta_p = \frac{u\sqrt{3}}{1 \pm 3v} \qquad (0 < \theta_p < \pi),$$

and $u$ and $v$ are the unique pair of positive integers for which $p = 3u^2 + (1 \pm 3v)^2$; and that $\psi_4(n) = 0$ for other values.

In the case in which $\alpha = 8$ I find that, if

$$n = (2^{a_2}.5^{a_5}.11^{a_{11}}\dots)^2.7^{a_7}.13^{a_{13}}.19^{a_{19}}\dots,$$

where 2, 5, 11, ... are the primes of the form $3m-1$ and 7, 13, 19, ... are those of the form $6m+1$, then

$$\frac{\psi_8(n)}{n\sqrt{n}} = (-1)^{a_2+a_5+a_{11}+\dots} \frac{\sin 3(1+a_7)\theta_7}{\sin 3\theta_7} \cdot \frac{\sin 3(1+a_{13})\theta_{13}}{\sin 3\theta_{13}} \dots \quad \text{.........(128)},$$

where $\theta_p$ is the same as in (127); and that $\psi_8(n) = 0$ for other values.

The case in which $\alpha = 12$ will be considered in § 28.

In short, such evidence as I have been able to find, while not conclusive, points to the truth of the results conjectured in § 18.

24. Analysis similar to that of the preceding sections may be applied to some interesting arithmetical functions of a different kind. Let

$$\phi^s(q) = 1 + 2\sum_1^\infty r_s(n) q^n \quad \text{..............................(129)},$$

where

$$\phi(q) = 1 + 2q + 2q^4 + 2q^9 + \dots,$$

so that $r_s(n)$ is the number of representations of $n$ as the sum of $s$ squares. Further let

$$\sum_1^\infty \delta_2(n) q^n = 2\left(\frac{q}{1-q} - \frac{q^3}{1-q^3} + \frac{q^5}{1-q^5} - \dots\right) = 2\left(\frac{q}{1+q^2} + \frac{q^2}{1+q^4} + \frac{q^3}{1+q^6} + \dots\right) \quad \text{......(130)};$$

$$(2^s - 1) B_s \sum_1^\infty \delta_{2s}(n) q^n = s\left(\frac{1^{s-1}q}{1+q} + \frac{2^{s-1}q^2}{1-q^2} + \frac{3^{s-1}q^3}{1+q^3} + \dots\right) \quad \text{.....................(131)},$$

when $s$ is a multiple of 4;

$$(2^s - 1) B_s \sum_1^\infty \delta_{2s}(n) q^n = s\left(\frac{1^{s-1}q}{1-q} + \frac{2^{s-1}q^2}{1+q^2} + \frac{3^{s-1}q^3}{1-q^3} + \dots\right) \quad \text{..............(132)},$$

when $s+2$ is a multiple of 4;

$$E_s \sum_1^\infty \delta_{2s}(n) q^n = 2^s\left(\frac{1^{s-1}q}{1+q^2} + \frac{2^{s-1}q^2}{1+q^4} + \frac{3^{s-1}q^3}{1+q^6} + \dots\right) + 2\left(\frac{1^{s-1}q}{1-q} - \frac{3^{s-1}q^3}{1-q^3} + \frac{5^{s-1}q^5}{1-q^5} - \dots\right) \quad \text{......(133)},$$

when $s-1$ is a multiple of 4;

$$E_s \sum_1^\infty \delta_{2s}(n) q^n = 2^s \left(\frac{1^{s-1} q}{1+q^2} + \frac{2^{s-1} q^2}{1+q^4} + \frac{3^{s-1} q^3}{1+q^6} + \dots\right) - 2\left(\frac{1^{s-1} q}{1-q} - \frac{3^{s-1} q^3}{1-q^3} + \frac{5^{s-1} q^5}{1-q^5} - \dots\right) \quad \dots\dots(134),$$

when $s+1$ is a multiple of 4. In these formulae

$$B_2 = \tfrac{1}{6}, \quad B_4 = \tfrac{1}{30}, \quad B_6 = \tfrac{1}{42}, \quad B_8 = \tfrac{1}{30}, \quad B_{10} = \tfrac{5}{66}, \dots$$

are Bernoulli's numbers, and

$$E_1 = 1, \quad E_3 = 1, \quad E_5 = 5, \quad E_7 = 61, \quad E_9 = 1385, \dots$$

are Euler's numbers. Then $\delta_{2s}(n)$ is in all cases an arithmetical function depending on the real divisors of $n$; thus, for example, when $s+2$ is a multiple of 4, we have

$$(2^s - 1) B_s \delta_{2s}(n) = s\{\sigma_{s-1}(n) - 2^s \sigma_{s-1}(\tfrac{1}{4}n)\} \quad \dots\dots(135),$$

where $\sigma_s(x)$ should be considered as equal to zero if $x$ is not an integer.

Now let

$$r_{2s}(n) = \delta_{2s}(n) + e_{2s}(n) \quad \dots\dots(136).$$

Then I can prove (see § 26) that

$$e_{2s}(n) = 0 \quad \dots\dots(137)$$

if $s = 1, 2, 3, 4$; and that

$$e_{2s}(n) = O(n^{s-1-\frac{1}{2}[\frac{2}{3}s]+\epsilon}) \quad \dots\dots(138)$$

for all positive integral values of $s$. But it is easy to see that, if $s \geqslant 3$, then

$$Hn^{s-1} < \delta_{2s}(n) < Kn^{s-1} \quad \dots\dots(139),$$

where $H$ and $K$ are positive constants. It follows that

$$r_{2s}(n) \sim \delta_{2s}(n) \quad \dots\dots(140)$$

for all positive integral values of $s$.

It appears probable, from the empirical results I obtain at the end of this paper, that

$$e_{2s}(n) = O\{n^{\frac{1}{2}(s-1)+\epsilon}\} \quad \dots\dots(141)$$

for all positive integral values of $s$; and that

$$e_{2s}(n) \neq o\{n^{\frac{1}{2}(s-1)}\} \quad \dots\dots(142)$$

if $s \geqslant 5$. But all that I can actually prove is that

$$e_{2s}(n) = O(n^{s-1-\frac{1}{2}[\frac{2}{3}s]}) \quad \dots\dots(143)$$

if $s \geqslant 9$; and that

$$e_{2s}(n) \neq o(n^{\frac{1}{2}s-1}) \quad \dots\dots(144)$$

if $s \geqslant 5$.

25. Let

$$f_{2s}(q) = \sum_1^\infty e_{2s}(n) q^n = \sum_1^\infty \{r_{2s}(n) - \delta_{2s}(n)\} q^n \quad \dots\dots(145).$$

Then it can be shown by the theory of elliptic functions that

$$f_{2s}(q) = \phi^{2s}(q) \sum_{1 \leqslant n \leqslant \frac{1}{4}(s-1)} K_n (kk')^{2n} \quad \dots\dots(146),$$

that is to say that

$$f_{2s}(q) = \frac{f^{4s}(-q)}{f^{2s}(q^2)} \sum_{1 \leqslant n \leqslant \frac{1}{4}(s-1)} K_n \frac{f^{24n}(q^2)}{f^{24n}(-q)} \quad \text{.......................(147)},$$

where $\phi(q)$ and $f(q)$ are the same as in § 13. We thus obtain the results contained in the following table.

TABLE VI.

1. $f_2(q) = 0, \quad f_4(q) = 0, \quad f_6(q) = 0, \quad f_8(q) = 0.$
2. $5f_{10}(q) = 16\dfrac{f^{14}(q^2)}{f^4(-q)}, \quad f_{12}(q) = 8f^{12}(q^2).$
3. $61f_{14}(q) = 728f^4(-q)f^{10}(q^2), \quad 17f_{16}(q) = 256f^8(-q)f^8(q^2).$
4. $1385f_{18}(q) = 24416f^{12}(-q)f^6(q^2) - 256\dfrac{f^{30}(q^2)}{f^{12}(-q)}.$
5. $31f_{20}(q) = 616f^{16}(-q)f^4(q^2) - 128\dfrac{f^{28}(q^2)}{f^8(-q)}.$
6. $50521f_{22}(q) = 1103272f^{20}(-q)f^2(q^2) - 821888\dfrac{f^{26}(q^2)}{f^4(-q)}.$
7. $691f_{24}(q) = 16576f^{24}(-q) - 32768f^{24}(q^2).$

It follows from the last formula of Table VI that

$$\tfrac{691}{64}e_{24}(n) = (-1)^{n-1}259\tau(n) - 512\tau(\tfrac{1}{2}n) \quad \text{.....................(148)},$$

where $\tau(n)$ is the same as in § 16, and $\tau(x)$ should be considered as equal to zero if $x$ is not an integer.

Results equivalent to 1, 2, 3, 4 of Table VI were given by Dr Glaisher in the *Quarterly Journal*, Vol. XXXVIII. The arithmetical functions called by him

$$\chi_4(n), \quad \Omega(n), \quad W(n), \quad \Theta(n), \quad U(n)$$

are the coefficients of $q^n$ in

$$\frac{f^{14}(q^2)}{f^4(-q)}, \quad f^{12}(q^2), \quad f^4(-q)f^{10}(q^2), \quad f^8(q)f^8(q^2), \quad f^{12}(-q)f^6(q^2).$$

He gave reduction formulae for these functions and observed how the functions which I call $e_{10}(n)$, $e_{12}(n)$ and $e_{16}(n)$ can be defined by means of the complex divisors of $n$. It is very likely that $\tau(n)$ is also capable of such a definition.

26. Now let us consider the order of $e_{2s}(n)$. It is easy to see from (147) that $f_{2s}(q)$ can be expressed in the form

$$\Sigma K_{a,b,c,h,k}\{f^3(-q)\}^a \left\{\frac{f^5(-q)}{f^2(q^2)}\right\}^b \left\{\frac{f^5(q^2)}{f^2(-q)}\right\}^c f^h(-q)f^k(q^2) \quad \text{...........(149)},$$

where $a$, $b$, $c$, $h$, $k$ are zero or positive integers, such that

$$a+b+c = [\tfrac{2}{3}s], \quad h+k = 2s - 3[\tfrac{2}{3}s].$$

Proceeding as in § 13 we can easily show that

$$n^{-\frac{1}{2}[\frac{2}{3}s]}\, e_{2s}(n)$$

cannot be of higher order than the coefficient of $q^{24n}$ in

$$\phi^A(q)\phi^B(q^3)\phi^C(q^2) \qquad \text{.........(150)},$$

where $C$ is 0 or 1 and $A+B+C = 2s - 2\left[\frac{2}{3}s\right]$.

Now, if $s \geqslant 5$, $A+B+C \geqslant 4$; and so $A+B \geqslant 3$. Hence one at least of $A$ and $B$ is greater than 1. But we know that

$$\phi^2(q) = \Sigma O(\nu^\epsilon)\, q^\nu.$$

It follows that the coefficient of $q^{24n}$ in (150) is of order not exceeding

$$n^{\frac{1}{2}(A+B+C)-1+\epsilon}.$$

Thus

$$e_{2s}(n) = O\left(n^{s-1-\frac{1}{2}[\frac{2}{3}s]+\epsilon}\right) \qquad \text{.........(151)}$$

for all positive integral values of $s$.

27. When $s \geqslant 9$ we can obtain a slightly more precise result.

If $s \geqslant 16$ we have $A+B+C \geqslant 12$; and so $A+B \geqslant 11$. Hence one at least of $A$ and $B$ is greater than 5. But

$$\phi^6(q) = \Sigma O(\nu^2)\, q^\nu.$$

It follows that the coefficient of $q^{24n}$ in (150) is of order not exceeding

$$n^{\frac{1}{2}(A+B+C)-1},$$

or that

$$e_{2s}(n) = O\left(n^{s-1-\frac{1}{2}[\frac{2}{3}s]}\right) \qquad \text{.........(152)},$$

if $s \geqslant 16$. We can easily show that (152) is true when $9 \leqslant s < 16$ considering all the cases separately, using the identities

$$f^{12}(-q)f^6(q^2) = \{f^3(-q)\}^4\{f^3(q^2)\}^2,$$

$$\frac{f^{30}(q^2)}{f^{12}(-q)} = \left\{\frac{f^5(q^2)}{f^2(-q)}\right\}^6,$$

$$f^{16}(-q)f^4(q^2) = \left\{\frac{f^5(-q)}{f^2(q^2)}\right\}^4\left\{\frac{f^5(q^2)}{f^2(-q)}\right\}^2 f^2(q^2),$$

$$\frac{f^{28}(q^2)}{f^8(-q)} = \left\{\frac{f^5(q^2)}{f^2(-q)}\right\}^4\{f^3(q^2)\}^2 f^2(q^2), \ \ldots,$$

and proceeding as in the previous two sections.

The argument of §§ 14—15 may also be applied to the function $e_{2s}(n)$. We find that

$$e_{2s}(n) \neq o\left(n^{\frac{1}{2}s-1}\right) \qquad \text{.........(153)}$$

I leave the proof to the reader.

28. There is reason to suppose that

$$\left.\begin{aligned} e_{2s}(n) &= O\{n^{\frac{1}{2}(s-1+\epsilon)}\} \\ e_{2s}(n) &\neq o\{n^{\frac{1}{2}(s-1)}\} \end{aligned}\right\} \qquad (154),$$

if $s \geqslant 5$. I find, for example, that

$$\sum_1^\infty \frac{e_{10}(n)}{n^s} = \frac{e_{10}(1)}{1+2^{2-s}} \Pi_1 \Pi_2 \qquad (155),$$

where $$\Pi_1 = \frac{1}{(1-3^{4-2s})(1-7^{4-2s})(1-11^{4-2s})\ldots},$$

3, 7, 11, ... being the primes of the form $4n-1$, and

$$\Pi_2 = \frac{1}{(1-2c_5 \cdot 5^{-s} + 5^{4-2s})(1-2c_{13} \cdot 13^{-s} + 13^{4-2s})\ldots},$$

5, 13, 17, ... being the primes of the form $4n+1$, and

$$c_p = u^2 - (4v)^2,$$

where $u$ and $v$ are the unique pair of positive integers satisfying the equation $u^2 + (4v)^2 = p^2$.

The equation (155) is equivalent to the assertion that, if

$$n = (3^{a_3} \cdot 7^{a_7} \cdot 11^{a_{11}} \ldots)^2 \cdot 2^{a_2} \cdot 5^{a_5} \cdot 13^{a_{13}} \ldots,$$

where $a_p$ is zero or a positive integer, then

$$\frac{e_{10}(n)}{n^2 e_{10}(1)} = (-1)^{a_2} \frac{\sin 4(1+a_5)\theta_5}{\sin 4\theta_5} \cdot \frac{\sin 4(1+a_{13})\theta_{13}}{\sin 4\theta_{13}} \ldots \qquad (156),$$

where $$\tan \theta_p = \frac{u}{v} \qquad (0 < \theta_p < \tfrac{1}{2}\pi),$$

$u$ and $v$ being integers satisfying the equation $u^2 + v^2 = p$; and $e_{10}(n) = 0$ otherwise. If this is true then we should have

$$\left|\frac{e_{10}(n)}{e_{10}(1)}\right| \leqslant n^2 d(n) \qquad (157)$$

for all values of $n$ and

$$\frac{e_{10}(n)}{e_{10}(1)} \geqslant n^2 \qquad (158)$$

for an infinity of values of $n$. In this case we can prove that, if $n$ is the square of a prime of the form $4m-1$, then

$$\frac{e_{10}(n)}{e_{10}(1)} = n^2.$$

Similarly I find that

$$\sum_1^\infty \frac{e_{12}(n)}{n^s} = e_{12}(1) \prod_p \left(\frac{1}{1+2c_p \cdot p^{-s} + p^{5-2s}}\right) \qquad (159),$$

$p$ being an odd prime and $c_p^2 \leqslant p^5$. From this it would follow that

$$\left|\frac{e_{12}(n)}{e_{12}(1)}\right| \leqslant n^{\frac{5}{2}} d(n) \qquad (160)$$

for all values of $n$ and

$$\frac{e_{12}(n)}{e_{12}(1)} \geqslant n^{\frac{5}{2}} \qquad (161)$$

for an infinity of values of $n$.

Finally I find that

$$\sum_1^\infty \frac{e_{16}(n)}{n^s} = \frac{e_{16}(1)}{1+2^{3-s}} \prod_p \left(\frac{1}{1+2c_p \cdot p^{-s}+p^{7-2s}}\right) \quad \text{.....................(162)},$$

$p$ being an odd prime and $c_p^2 \leqslant p^7$. From this it would follow that

$$\frac{e_{16}(n)}{e_{16}(1)} \leqslant n^{\frac{7}{2}} d(n) \quad \text{.......................................(163)}$$

for all values of $n$ and

$$\left|\frac{e_{16}(n)}{e_{16}(1)}\right| \geqslant n^{\frac{7}{2}} \quad \text{...........................................(164)}$$

for an infinity of values of $n$.

In the case in which $2s = 24$ we have

$$\tfrac{691}{64} e_{24}(n) = (-1)^{n-1}\, 259\, \tau(n) - 512\, \tau(\tfrac{1}{2}n).$$

I have already stated the reasons for supposing that

$$|\tau(n)| \leqslant n^{\frac{11}{2}} d(n)$$

for all values of $n$ and

$$|\tau(n)| \geqslant n^{\frac{11}{2}}$$

for an infinity of values of $n$.

# X. *On the Fifth Book of Euclid's Elements* (Fourth Paper)*.

By M. J. M. Hill, M.A., LL.D., Sc.D., F.R.S.,
Astor Professor of Mathematics in the University of London.

[*Received* 11 December 1916. *Read* 5 February 1917.]

1. The object of this brief paper is to continue the discussion of the famous fifth definition in Arts. 4—8 of my third paper*.

In Art. 5 of that paper I stated that I believed that Stolz in his *Vorlesungen über allgemeine Arithmetik*, part I. p. 87, published in 1885, was the first to reduce to two the number of independent sets of conditions comprised in the fifth definition.

2. My attention has been recently called by my friend Mr Rose-Innes to a passage in De Morgan's article on Proportion in vol. XIX. of the *Penny Cyclopaedia*, published in 1841, from which it appears that the possibility of this reduction was known to him. It will be seen from the foot-note he appends to his demonstration that he was aware that his demonstration was not exact in form. The words "of given nearness" which he uses are difficult to interpret. I have however endeavoured to complete the proof on what I suppose were his lines of reasoning, or in the event of my having misinterpreted his words, then on the lines which his argument has suggested to me.

He says (*l.c.* p. 52, column 2):

> "It is however perfectly allowable to leave out of sight the possible case in which a multiple of $A$ is exactly equal to a multiple of $B$; since if the test be true in all other cases, it is therefore true in this. For, if possible, let $4A = 7B$, and $4C$ be (say) greater than $7D$. Then $m(4C)$ exceeds $m(7D)$ by $m$ times this difference, which may be made as great as we please, or $4mC$ and multiples succeeding it, may be made to fall in an interval as many intervals removed from that of $7mD$ and $(7m+1)D$ as we please. But $4mA$ is equal to $7mB$, whence $(4m+1)A$, &c. must fall among the multiples of $B$ in intervals of given nearness† to the interval of $7mB$ and $(7m+1)B$. Consequently the multiples of $A$ following $4mA$ cannot always fall among the multiples of $B$ in the same intervals as the same multiples of $C$ among those of $D$; and the rest of the test cannot be true unless $4C = 7D$; that is, if the rest of the test be true, then $4C = 7D$."

* The first paper will be found in vol. XVI, part IV, the second in vol. XIX, part II, and the third in vol. XXII.

† (De Morgan's foot-note.) We leave the reader to put this demonstration into a more exact form.

3. To prove that $4C$ cannot exceed $7D$, it is sufficient to show that if $4C$ do exceed $7D$, then a *single* pair of integers $p, q$ must exist such that whilst $pA < qB$, $pC > qD$. But De Morgan's idea seems* to be to find a number $n$ such that if $t$ be *any* positive integer *whatever* or zero then another positive integer $w$ exists such that whilst

$$(n+t)A < wB,$$

$$(n+t)C > wD.$$

4. I proceed to complete the demonstration on these lines.

It is supposed that, if $r$, $s$ be any two relatively prime integers except the pair $r_1$, $s_1$, it is given that

if $$rA > sB, \text{ then } rC > sD;$$

but if $$rA < sB, \text{ then } rC < sD.$$

Further it is given that $r_1A = s_1B$, and it is required to show that $r_1C = s_1D$.

(i) Suppose if possible that

$$r_1A = s_1B, \text{ but } r_1C > s_1D.$$

Then I imagine De Morgan's first step was to take an integer $m$ such that

$$m(r_1C - s_1D) > pD,$$

where $p$ is any selected integer however large.

In order to reach the result set out in Art. 3 it is enough to take $p = 1$.

Let us take for $m$ the smallest integer which makes

$$m(r_1C - s_1D) > D.$$

Next since $r_1A = s_1B$,

$$(mr_1 + t)A = \left(ms_1 + \frac{s_1t}{r_1}\right)B.$$

Let $w$ be the integer next greater than $ms_1 + \frac{s_1t}{r_1}$;

$$\therefore ms_1 + \frac{s_1t}{r_1} < w \leqq ms_1 + \frac{s_1t}{r_1} + 1,$$

and $$(mr_1 + t)A < wB.$$

* I have drawn this conclusion with some hesitation from De Morgan's use of the words "$4mC$ and multiples succeeding it" in the 5th and 6th lines of the extract quoted above, but it is not in accord with the words "cannot always" in the 9th and 10th lines, which suggest that he was desirous of establishing the existence of a difference in the distribution of the multiples of $A$ amongst those of $B$ as compared with the distribution of the multiples of $C$ amongst those of $D$ in some cases only, and that he was not aiming at establishing the existence of such a difference in the case of all multiples of $A$ following a certain multiple of $A$.

Further $$r_1C = s_1D + (r_1C - s_1D),$$

$$\therefore\ (mr_1 + t)C = \left(ms_1 + \frac{s_1t}{r_1}\right)D + \left(m + \frac{t}{r_1}\right)(r_1C - s_1D)$$

$$> \left(ms_1 + \frac{s_1t}{r_1}\right)D + m(r_1C - s_1D)$$

$$> \left(ms_1 + \frac{s_1t}{r_1}\right)D + D;$$

$$\therefore\ (mr_1 + t)C > wD,$$

but $$(mr_1 + t)A < wB.$$

This is true when $t$ is *any* positive integer *whatever* or zero.

Hence the multiples of $A$, from and after $mr_1A$, are not distributed amongst the multiples of $B$ in the same way as the multiples of $C$, from and after $mr_1C$, are distributed amongst the multiples of $D$.

[Since $$\left(ms_1 + \frac{s_1t}{r_1}\right) < w \overline{\gtrless} \left(ms_1 + \frac{s_1t}{r_1} + 1\right),$$

$$\therefore\ \frac{s_1}{r_1} < \frac{w}{mr_1 + t} \overline{\gtrless} \frac{s_1}{r_1} + \frac{1}{mr_1 + t}.$$

Hence as $t$ tends to $\infty$, $\frac{w}{mr_1 + t}$ tends to $\frac{s_1}{r_1}$.]

(ii) Suppose next that $r_1A = s_1B$, but $r_1C < s_1D$.

Let $m$ be the smallest integer which makes

$$m(s_1D - r_1C) > D;$$

$$\because\ r_1A = s_1B,$$

$$\therefore\ (mr_1 + t)A = \left(ms_1 + \frac{s_1t}{r_1}\right)B.$$

Let $w$ be the integer next below $ms_1 + \frac{s_1t}{r_1}$;

$$\therefore\ ms_1 + \frac{s_1t}{r_1} - 1 \leqq w < ms_1 + \frac{s_1t}{r_1},$$

and $$(mr_1 + t)A > wB.$$

Now $$r_1C + (s_1D - r_1C) = s_1D,$$

$$\therefore\ (mr_1 + t)C + \left(m + \frac{t}{r_1}\right)(s_1D - r_1C) = \left(ms_1 + \frac{s_1t}{r_1}\right)D;$$

$$\therefore\ (mr_1 + t)C + m(s_1D - r_1C) < \left(ms_1 + \frac{s_1t}{r_1}\right)D;$$

$$\therefore\ (mr_1 + t)C + D < \left(ms_1 + \frac{s_1t}{r_1}\right)D;$$

$$\therefore\ (mr_1 + t)C < wD,$$

but $$(mr_1 + t)A > wB.$$

This is true when $t$ is *any* positive integer *whatever* or zero.

Hence the multiples of $A$, from and after $mr_1A$, are not distributed amongst the multiples of $B$ in the same way as the multiples of $C$, from and after $mr_1C$, are distributed amongst the multiples of $D$.

[Since
$$ms_1 + \frac{s_1 t}{r_1} - 1 \leqq w < ms_1 + \frac{s_1 t}{r_1},$$
$$\therefore \frac{s_1}{r_1} - \frac{1}{mr_1 + t} \leqq \frac{w}{mr_1 + t} < \frac{s_1}{r_1}.$$

Hence as $t$ tends to $\infty$, $\frac{w}{mr_1 + t}$ tends to $\frac{s_1}{r_1}$.]

5. I proceed next to give Stolz's proof.

He states the theorem thus:

Suppose that for every pair $r$, $s$ of relatively prime numbers, excepting a single pair $r_1$, $s_1$, there correspond to the relations $rA \gtreqless sB$ the relations $rC \gtreqless sD$, whilst $r_1A = s_1B$, then must $r_1C = s_1D$.

Let $m$ be any positive whole number, not divisible by $r_1$, such that $mA > B$, then $mA$ cannot be a multiple of $B$. Consequently an integer $n$ exists such that
$$nB < mA < (n+1)B;$$
$$\therefore nr_1B < mr_1A < (n+1)r_1B;$$
$$\therefore nr_1B < ms_1B < (n+1)r_1B;$$
$$\therefore nr_1 < ms_1 < (n+1)r_1.$$

But since
$$nB < mA < (n+1)B,$$
therefore by hypothesis,
$$nD < mC < (n+1)D.$$

If possible, let $r_1C \neq s_1D$.

(i) Suppose first that
$$C > \frac{s_1}{r_1}D.$$

Then since
$$\frac{s_1}{r_1} > \frac{n}{m} \text{ and } C < \frac{n+1}{m}D,$$
$$\therefore C - \frac{s_1}{r_1}D < \frac{n+1}{m}D - \frac{n}{m}D;$$
$$\therefore C - \frac{s_1}{r_1}D < \frac{D}{m}.$$

*But this cannot be, because $m$ can be taken so large that $\frac{D}{m}$ is less than any quantity of the same kind as $D$, and therefore less than $C - \frac{s_1}{r_1}D$;
$$\therefore C \ngtr \frac{s_1}{r_1}D.$$

* This is the starting point of the proof given in the next article, where however $m$ has a different meaning.

(ii) Next, if $C < \frac{s_1}{r_1} D$,

then since $$\frac{s_1}{r_1} < \frac{n+1}{m} \text{ and } C > \frac{n}{m} D,$$

$$\therefore \frac{s_1}{r_1} D - C < \frac{n+1}{m} D - \frac{n}{m} D:$$

$$\therefore \frac{s_1}{r_1} D - C < \frac{D}{m}$$

which is impossible, as before.

$$\therefore C \nless \frac{s_1}{r_1} D.$$

Consequently, $$C = \frac{s_1}{r_1} D,$$

$$\therefore r_1 C = s_1 D.$$

6. The following proof, which I gave in my Second Paper, Art. 67, was sent me by both Prof. A. C. Dixon and Mr E. Budden. It is, in fact, Stolz's proof reversed, but it seems simpler and more direct.

The enunciation is of course the same as Stolz's.

(i) Suppose $r_1 A = s_1 B$, but $r_1 C > s_1 D$.

Let $m$ be the smallest integer which makes

$$m(r_1 C - s_1 D) > D.$$

$$\therefore mr_1 C > (ms_1 + 1) D,$$

but $$mr_1 A = ms_1 B;$$

$$\therefore mr_1 A < (ms_1 + 1) B,$$

whilst $$mr_1 C > (ms_1 + 1) D,$$

which is inconsistent with the condition that all values of $r$, $s$ which make $rA < sB$ must also make $rC < sD$. Therefore $r_1 C \ngtr s_1 D$.

(ii) Suppose next $r_1 A = s_1 B$, but $r_1 C < s_1 D$.

Now let $m$ be the smallest integer which makes

$$m(s_1 D - r_1 C) > C,$$

$$\therefore (mr_1 + 1) C < ms_1 D,$$

but $$mr_1 A = ms_1 B;$$

$$\therefore (mr_1 + 1) A > ms_1 B,$$

but $$(mr_1 + 1) C < ms_1 D.$$

This is inconsistent with the condition that all values of $r$, $s$ which make $rA > sB$ must also make $rC > sD$. Therefore $r_1 C \nless s_1 D$.

Consequently, $$r_1 C = s_1 D.$$

7. It appears that the possibility of the reduction to two of the number of sets of conditions in the fifth definition was first enunciated by De Morgan, that the first clear and unambiguous proof is due to Stolz, but the simplest and most direct proof is that due to Prof. Dixon and Mr Budden.

# XI. *The Character of the Kinetic Potential in Electromagnetics.* By R. HARGREAVES, M.A.

[*Received* 14 November 1916. *Read* 5 February 1917.]

THERE are three important volume integrals over an infinite electromagnetic field derivable from Maxwell's equations as modified by FitzGerald and Lorentz. One deals with flux of energy, another with flux of momentum, while the third gives an expression for the difference between electrical and magnetic energies. This last quantity has been called the kinetic potential, and the term carries with it the suggestion of an advance from the electromagnetic stage in which an infinite field is considered, to a dynamical problem in which the activities of the field are summed up in a single expression dependent on the coordinates of charges, and on their derivatives with respect to time. A normal kinetic potential constitutes in fact the complete statement of a dynamical problem.

There is one *primâ facie* reason for doubting the normal character of this function in Electromagnetics, viz. the necessity of accounting for dissipation of energy and momentum. It is the first object of this paper to establish a departure from normal character, and to fix its precise nature. As dissipation is due to flux at infinity, which is a feature of the other integrals cited above, it is plain that if the kinetic potential fails in any of its normal functions we should look to these integrals to fill the lacuna. It is our second object to shew that they are adequate for the purpose, and to deal with the method of applying them.

§ 1. Of the two main purposes which the K.P. serves, the derivation of expressions for energy and force, the former is that in which help is specially needed. Sommerfeld in his memoir on potentials calls attention to the difficulty (no doubt experienced by others) of dealing with the energy, on the ground that a quadratic function of electric and magnetic vectors is to be integrated. But the K.P., which is also such a function, *is* integrable so far as its effective part is concerned, when the potentials are known. If the method of K.P. had complete validity it would be possible to turn the difficulty as regards energy by deriving it from the K.P. The form obtained for the K.P. is

$$\Sigma \frac{e}{2V}(Fu + Gv + Hw - \psi V) + \frac{d\chi}{dt},$$

in which $(uvw)$ is the velocity of an element $e$ of charge, $V$ that of propagation, $\psi$, $F$, $G$, $H$ are potentials, and $\chi$ is an integral not evaluated. For use as a K.P. a function which is a complete time-rate is entirely nugatory. The sum

$$\Sigma \frac{e}{2}(Fu + Gv + Hw - \psi V)$$

in the case of a finite number of point-charges is a completely integrated form when the values of the potentials are known, and the way is therefore clear for the application of the method to point-charges. This then is a case in which we can make the enquiry as to the character of the K.P., and the proper use of the other integrals. Point-action has always been the basis of mathematical treatment for continuous distributions, and evidently has a more immediate application in an atomic electrical theory.

With respect to the other function of a K.P., the derivation of force, there is no such disability in the direct application of electromagnetic methods. The connexion between the methods may be briefly stated. If $\psi_1$, ... are values due to the action of $e_1$, estimated at a place $(x_2 y_2 z_2)$, where at time $t$ there is a charge $e_2$ moving with velocity $(u_2 v_2 w_2)$, the force on $e_2$ due to $e_1$ has the $x$ component

$$\begin{aligned} e_2\{X_1 + (v_2 c_1 - w_2 b_1)\, V\} &= -e_2\left(\frac{1}{V}\frac{\partial F_1}{\partial t} + \frac{\partial \psi_1}{\partial x_2}\right) + \frac{e_2 v_2}{V}\left(\frac{\partial G_1}{\partial x_2} - \frac{\partial F_1}{\partial y_2}\right) - \frac{e_2 w_2}{V}\left(\frac{\partial F_1}{\partial z_2} - \frac{\partial H_1}{\partial x_2}\right) \\ &= -\frac{e_2}{V}\left(\frac{\partial F_1}{\partial t} + u_2\frac{\partial F_1}{\partial x_2} + v_2\frac{\partial F_1}{\partial y_2} + w_2\frac{\partial F_1}{\partial z_2}\right) + \frac{e_2}{V}\frac{\partial}{\partial x_2}(F_1 u_2 + G_1 v_2 + H_1 w_2 - \psi_1 V), \end{aligned}$$

since for a point-charge $\frac{\partial u_2}{\partial x_2}$... do not exist. Now $\frac{\partial F_1}{\partial t}$ includes the whole dependence of $F_1$ on time through the motion of $e_1$, and therefore

$$\frac{\partial F_1}{\partial t} + u_2\frac{\partial F_1}{\partial x_2} + v_2\frac{\partial F_1}{\partial y_2} + w_2\frac{\partial F_1}{\partial z_2}$$

is a complete time-rate $\frac{dF_1}{dt}$. Thus if we write

$$L = \frac{e_2}{V}(F_1 u_2 + G_1 v_2 + H_1 w_2 - \psi_1 V), \qquad \frac{\partial L}{\partial u_2} = e_2\frac{F_1}{V},$$

and the above expression has the form

$$\frac{\partial L}{\partial x_2} - \frac{d}{dt}\frac{\partial L}{\partial u_2}.$$

We observe at once that this is a derivation, not from the kinetic potential

$$\frac{e_2}{2V}(F_1 u_2 + G_1 v_2 + H_1 w_2 - \psi_1 V) + \frac{e_1}{2V}(F_2 u_1 + G_2 v_1 + H_2 w_1 - \psi_2 V),$$

but from the first section doubled. This is a first hint of departure from normal conditions. In the statical problem there is equality of the two sections; in the general problem dynamical equivalence is consistent with inequality if the difference is a complete differential coefficient with respect to the time, or more briefly if the difference takes the form $\frac{df}{dt}$, so that there is no immediate certainty of breach of normal conditions.

§ 2. To make a further step the formula for the potential of a point-charge is expanded in a series proceeding by inverse powers of $V$, the velocity of propagation. This value is required only at the point where the second charge is placed, and is used to give a corresponding expansion for the K.P. The groups containing odd and even powers of $V$ are considered separately. It then appears that the terms of even order in the two parts named above do in fact differ by a quantity of the form $\frac{df}{dt}$. The even sections of the two parts are therefore equivalent, and their sum may be accounted a true kinetic potential, which if it stood alone would represent a system conservative in respect to energy and momentum linear or angular.

The terms of odd order possess a property the antithesis of that of the conservative group, viz. the *sum* of these terms in the two parts has the form $\frac{df}{dt}$. Thus, if we write $L_{12}$ and $L'_{12}$ for even and odd sections of one part, we find that $L_{12} - L_{21}$ has the form $\frac{df}{dt}$, while $L'_{12} + L'_{21}$ has that form. The force of even order on $e_2$ given by $2L_{12}$ is also given by $2L_{21}$ or by $L_{12} + L_{21}$; but the force of odd order given by $2L'_{12}$ is also given by $-2L'_{21}$, while $L'_{12} + L'_{21}$ which we expect to be the K. P. in fact yields no force. There is then a distinct breach of normal conditions in the group of terms of odd order specially associated with radiation, which we may call the dissipative group.

§ 3. The volume integral involving rate of change of energy has the form

$$0 = \frac{dE}{dt} + \Sigma\,(\xi\dot{x} + \eta\dot{y} + \zeta\dot{z}) + \text{flux of energy (or total radiation)},$$

if $E$ is the total electromagnetic energy, $\xi$ the $x$ component of electromagnetic force on a point-charge with $\dot{x}$ for component of velocity. But if we use $E$ and $\xi$ for the sections of even order, $E'$ and $\xi'$ for those of odd order, we have

$$0 = \frac{dE}{dt} + \Sigma\,(\xi\dot{x} + \eta\dot{y} + \zeta\dot{z})$$

in virtue of derivation from a regular K.P.; and

$$0 = \frac{dE'}{dt} + \Sigma\,(\xi'\dot{x} + \eta'\dot{y} + \zeta'\dot{z}) + \text{flux of energy};$$

or in effect the flux of energy is a quantity of odd order in $V$. This flux is found by use of simplified values in an integral over an infinite sphere. When this is evaluated it should appear that *the sum of radiation and a rate of working is a complete time-rate*, and then an expression for $E'$ can be found. The equation for loss of energy is got by writing

$$\Sigma\,\{(\xi + \xi')\,\dot{x} + (\eta + \eta')\,\dot{y} + (\zeta + \zeta')\,\dot{z}\}, = \frac{dT}{dt} - \Sigma\,(\dot{x}f_x + \dot{y}f_y + \dot{z}f_z),$$

where $T$ is the material kinetic energy, and $f_x$ a component of mechanical force if any such exists. That equation is

$$\frac{d}{dt}(T + E + E') + \text{Radiation (or flux of energy)} = \Sigma\,(\dot{x}f_x + \dot{y}f_y + \dot{z}f_z).$$

We have first the use of a formal relation between different electromagnetic quantities to give an expression for $E'$, and later, when we associate electromagnetic with extraneous mechanical forces, or definitely postulate their absence, we have an equation for rate of loss of total energy.

The flux of momentum is also of odd order, and the second volume integral gives a formal relation of the type

$$0 = \xi' + \frac{dp'}{dt} + \text{flux of momentum},$$

which is connected with mechanical forces by writing $\xi + \xi' = \Sigma m\ddot{x} - f_x$.

Here again two quantities, now a flux and a force, are to give on summation a complete time-rate, and we infer the expression of a momentum $p'$ belonging to the dissipative group of terms. In the case of each integral a condition is to be satisfied which will give us the assurance that the flux at infinity is correctly treated. Also it will be found to involve the localizing of the parts of these integrals, a problem solved for the conservative section by the use of K. P.

§ 4. It is understood in Dynamics that all coordinates are stated with reference to one time. In Electromagnetics the primary position is that coordinates of a source are referred to the time of departure of a wave, and those of a point affected to a time of arrival. As a preliminary to dynamical treatment we require that coordinates and velocities of the source should be referred to the time of arrival. If $(x_2 y_2 z_2 t)$ denote place and time of arrival of a disturbance originating from $e_1$ a point-charge at place and time $(x_1' y_1' z_1' t')$, $x_1'$... being functions of $t'$, this reference is made by the use of

$$r'^2 = (x_2 - x_1')^2 + (y_2 - y_1')^2 + (z_2 - z_1')^2, \quad \text{and} \quad t' = t - r'/V \quad \text{.................(1)}.$$

The values of potentials at $(x_2 y_2 z_2 t)$ are

$$\psi_1 = e_1/\{r' - \Sigma (x_2 - x_1') u_1'/V\}, \quad F_1 = u_1' \psi_1/V \quad \text{.......................(2)};$$

and these values are to be expressed through (1) in terms of $x_2 - x_1$, $u_1$, $\dot{u}_1$..., where

$$x_1 = x_1(t), \text{ and } u_1 = \dot{x}_1(t), \ldots.$$

We use $r$ for $\sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2 + (z_2 - z_1)^2}$,

and $(xyz)$ or $(lr, mr, nr)$ for $(x_2 - x_1, \quad y_2 - y_1, \quad z_2 - z_1)$.

The values of potentials and of the resulting K. P., which would follow from a treatment of (2) by successive approximations, may be obtained more rapidly as follows. The expansion of $\psi_1$ is

$$\psi_1 = e_1 \left\{\frac{1}{r} + \frac{D_1^2 r}{V^2 2!} - \frac{D_1^3 r^2}{V^3 3!} + \frac{D_1^4 r^3}{V^4 4!} - \ldots\right\} \quad \text{.........................(3)},$$

where $D_1$ or $\frac{d_1}{dt}$ denotes differentiation through $x_1 y_1 z_1$ and their derivatives $u_1$, ....

If we operate with $\nabla^2$ on $\psi_1$, where $x_2 y_2 z_2$ are the space coordinates in $\nabla^2$, and use the equation $\nabla^2 r^n = n(n+1) r^{n-2}$, then noting that $\nabla^2$ and $D_1$ are commutative, we get

$$\nabla^2 \psi_1 = D_1^2 \frac{e_1}{V^2} \left\{\frac{1}{r} + \frac{D_1^2 r}{V^2 2!} - \frac{D_1^3 r^2}{V^3 3!} + \ldots\right\} = \frac{D_1^2 \psi_1}{V^2},$$

the fundamental equation for $\psi_1$. For the expansion of $F_1$ we have

$$F_1 = \frac{e_1}{V} \left\{\frac{u_1}{r} - \frac{D_1 u_1}{V} + \frac{D_1^2 r u_1}{V^2 2!} - \frac{D_1^3 r^2 u_1}{V^3 3!} + \ldots\right\} \quad \text{.......................(4)}.$$

It follows that the expansion of $L_{12}$, the part of the K. P. due to the potentials of $e_1$ at $(x_2y_2z_2)$ the place of $e_2$, is given by

$$2L_{12} = -\frac{e_1e_2}{V^2}\left\{\frac{V^2-\Sigma u_1u_2}{r} - \frac{D_1(V^2-\Sigma u_1u_2)}{V} + \frac{D_1{}^2 r(V^2-\Sigma u_1u_2)}{V^2\,2!} - \frac{D_1{}^3 r^2(V^2-\Sigma u_1u_2)}{V^3\,3!} + \dots\right\} \dots(5).$$

Consider separately the terms of even order in $V$ with a view to obtaining an equivalent form, i.e. one differing by a term $\frac{df}{dt}$, of symmetrical character. Now in (5) where a K.P. is in question $x_2$ as well as $x_1$ is a function of $t$, and we may use $D_2$ or $\frac{d_2}{dt}$ for differentiation through $x_2$ and its derivatives, so that $\frac{d}{dt} = D_1 + D_2$. A term $D_1{}^2 f$ may be replaced by $-D_1D_2f$, a term $D_1{}^4 f$ by $D_1{}^2D_2{}^2f$, and so on. Hence an equivalent of the even group in $2L_{12}$ is $L$ where

$$L = -\frac{e_1e_2}{V^2}\left[\frac{V^2-\Sigma u_1u_2}{r} - \frac{D_1D_2 r(V^2-\Sigma u_1u_2)}{V^2\,2!} + \frac{D_1{}^2D_2{}^2 r^3(V^2-\Sigma u_1u_2)}{V^4\,4!} - \dots\right] \dots\dots(6),$$

a form symmetrical with respect to the two points, and therefore replacing even terms of $2L_{12}$ for $e_2$ and of $2L_{21}$ for $e_1$.

For the group of odd terms, $D_1$ may be put outside the bracket, and even powers of $D_1$ within it replaced by $-D_1D_2$, $D_1{}^2D_2{}^2, \dots$, giving for the equivalent K.P.

$$2L'_{12} = D_1\frac{e_1e_2}{V^2}\left[\frac{V^2-\Sigma u_1u_2}{V} - \frac{D_1D_2 r^2(V^2-\Sigma u_1u_2)}{V^3\,3!} + \frac{D_1{}^2D_2{}^2 r^4(V^2-\Sigma u_1u_2)}{V^5\,5!} - \dots\right] \dots(7).$$

An equivalent of the odd terms in $2L_{21}$ has the same form with $D_2$ outside the bracket, and the sum has therefore the form $\frac{df}{dt}$. The points stated in § 2 are therefore established.

§ 5. As a formal example of these results we may apply them to the case of charges $e_1$ and $e_2$ distributed uniformly over the surfaces of concentric spheres of expanding radii $r_1$ and $r_2$. If we write $e_1d\omega_1/4\pi$ for $e_1$ and $e_2d\omega_2/4\pi$ for $e_2$ the integrand for joint terms (containing $e_1$ and $e_2$) contains direction cosines only in the combination $\Sigma l_1l_2$, and an element of integration $d\omega_1d\omega_2/(4\pi)^2$ may be replaced by $dn/2$, where $n = \Sigma l_1l_2$ and has limits $-1$ and $+1$.

Whether in (6) or in the bracket of (7), terms of like order in $V$ give an operator (which is a power of $D_1D_2$) acting on

$$D_1D_2(r_1{}^2+r_2{}^2-2nr_1r_2)^{\frac{m+1}{2}} + (m+1)(m+2)\,n\dot{r}_1\dot{r}_2(r_1{}^2+r_2{}^2-2nr_1r_2)^{\frac{m-1}{2}}.$$

This quantity is equal to $-\dot{r}_1\dot{r}_2(m+1)\frac{d}{dn}(1-n^2)(r_1{}^2+r_2{}^2-2nr_1r_2)^{\frac{m-1}{2}}$,

and the integral between limits $-1$ and $+1$ for $n$ vanishes. Another form of the relation is

$$\int_{-1}^{+1}\{(m+1)\,r_1r_2(1-n^2)\,R^{m-1} + 2nR^{m+1}\}\,dn = 0$$

with $R^2 = r_1{}^2 + r_2{}^2 - 2nr_1r_2$. Thus in (6) and in (7) the total is reduced to the term of lowest order in $V^{-1}$, which in (7) is a constant; i.e. there are no odd terms, and radiation is absent. The conservative K.P. is reduced to $-\frac{e_1 e_2}{2}\int_{-1}^{+1}\frac{dn}{R}$ or to $-\frac{e_1 e_2}{r_2}$ if $r_2 > r_1$. From this we infer at once the terms $-\frac{e_1^2}{2r_1}$ and $-\frac{e_2^2}{2r_2}$ due to the actions of elements of one sphere on other elements of the same sphere; and the K.P. is that due to a statical system.

This simplicity however cannot attach to the potentials, for the vector potentials do not vanish and so the scalar potential cannot be independent of the motion. We may examine these potentials for a single sphere, radius $r_1$, charge $e_1$. Symmetry justifies us in writing

$$F, G, H = \left(\frac{x}{r}, \frac{y}{r}, \frac{z}{r}\right)\times S,$$

which through $\frac{\partial G}{\partial x} - \frac{\partial F}{\partial y} = 0$ makes the magnetic vector vanish and so involves absence of radiation. The equation

$$\frac{\partial F}{\partial x} + \frac{\partial G}{\partial y} + \frac{\partial H}{\partial z} + \frac{1}{V}\frac{\partial \psi}{\partial t} = 0$$

is then represented by

$$\frac{\partial S}{\partial r} + \frac{2S}{r} + \frac{1}{V}\frac{\partial \psi}{\partial t} = 0 \quad\text{.................(8 a).}$$

As

$$-X = \frac{1}{V}\frac{\partial F}{\partial t} + \frac{\partial \psi}{\partial x} = \frac{x}{r}\left(\frac{1}{V}\frac{\partial S}{\partial t} + \frac{\partial \psi}{\partial r}\right),$$

electrical force is reduced to the statical value if

$$\frac{1}{V}\frac{\partial S}{\partial t} + \frac{\partial \psi}{\partial r} = -\frac{e}{r^2} \text{ or } 0 \quad\text{.................(8 b),}$$

according as $r \gtrless r_1$.

Now take $(0, 0, r)$ for the point at which we are seeking potentials, so that $F = 0$, $G = 0$, and $H = S$. By (3) and (4), with $R^2 = r^2 + r_1^2 - 2nr_1 r$, we have

$$\left.\begin{aligned} \psi &= \frac{e_1}{2}\int_{-1}^{+1}\left[\frac{1}{R} + \frac{D_1^2 R}{V^2 2!} - \frac{D_1^3 R^2}{V^3 3!} + \frac{D_1^4 R^3}{V^4 4!}\cdots\right]dn \\ S &= \frac{e_1}{2V}\int_{-1}^{+1}\left[\frac{\dot r_1}{R} - \frac{D_1 \dot r_1}{V} + \frac{D_1^2(R\dot r_1)}{V^2 2!} - \frac{D_1^3(R^2 \dot r_1)}{V^3 3!}\cdots\right]n\,dn \end{aligned}\right\}\quad\text{.................(9).}$$

The properties (8 a) and (8 b) then hold, if

$$\dot r_1 \frac{\partial}{\partial r}\int_{-1}^{+1} r^2 R^m n\,dn + D_1 \int_{-1}^{+1} r^2 R^m dn = 0$$

and

$$\frac{\partial}{\partial r} D_1 \int_{-1}^{+1}\frac{R^{m+2}dn}{(m+2)(m+3)} + \dot r_1 \int_{-1}^{+1} R^m n\,dn = 0.$$

The latter is obviously the result given above, and the former is the second form given to the same result above.

These properties enable us to write

$$S = \frac{\partial \phi}{\partial r}, \quad \psi + \frac{1}{V}\frac{\partial \phi}{\partial t} = \frac{e_1}{r} \text{ or } \frac{e_1}{r_1}$$

according as $r \gtrless r_1$: the quantity $\phi$ is equal to

$$-\frac{e_1}{2}\int_{-1}^{+1}\left[\frac{D_1 R}{V\,2!}-\frac{D_1^2 R^2}{V^2\,3!}+\frac{D_1^3 R^3}{V^3\,4!}\cdots\right]dn,$$

and $\nabla^2\phi - \frac{1}{V^2}\frac{\partial^2\phi}{\partial t^2} = 0$ or $\frac{e_1 \dot{r}_1}{r_1^2}$ according as $r \gtrless r_1$.

We have thus a satisfactory account of the properties of $\psi$ and $S$, in virtue of which complicated values of potentials due to the expanding motion are consistent with the system as a whole having the potential energy of a statical system.

§ 6. With a view to shewing that the method of partition into conservative and dissipative groups leads to reasonable results, we examine briefly the opening terms before proceeding to a general method of realizing the differentiations involved in the series. For the conservative group the efficient terms to the second order are

$$L = \frac{e_1 e_2}{r}\left(-1 + \frac{\Sigma u_1 u_2 + \Sigma l u_1 \Sigma l u_2}{2V^2}\right) \quad \text{..........................(10 } a),$$

and the corresponding terms in the energy are

$$E = \frac{e_1 e_2}{r}\left(1 + \frac{\Sigma u_1 u_2 + \Sigma l u_1 \Sigma l u_2}{2V^2}\right) \quad \text{..........................(10 } b).$$

The relation of (10 $a$) to the formula of Clausius may be noted. His first simplification of an original more general formula gave a value

$$U = \frac{e_1 e_2}{r}\{k\Sigma u_1 u_2 + k_1(\Sigma u_1 u_2 - \Sigma l u_1 \Sigma l u_2)\}.$$

With respect to $k_1$ he suggested that simplification in the force may be obtained either by writing $k_1 = 0$ or $k_1 = -k$. The first value is that which he preferred, and with that value the formula is known as Clausius's potential. The present form is obtained by taking $k_1 = -k/2$ and $k = 1/V^2$. It is the form properly belonging to Maxwell's theory, and here it is understood that in the application to moving charges it gives first and second order terms in a series of terms of even order, constituting the K.P. of the conservative section of the forces.

A component of force is given by

$$\xi_2 = \frac{\partial L}{\partial x_2} - \frac{d}{dt}\frac{\partial L}{\partial u_2} = \frac{e_1 e_2}{r^2}\left[l\left\{1 - \frac{1}{2V^2}(3\sigma_1^2 - \Sigma u_1^2 + 2\Sigma u_1 u_2 + r\dot{\sigma}_1)\right\} + \frac{u_1\sigma_2}{V^2} - \frac{r\dot{u}_1}{2V^2}\right]$$

in which $\sigma_1 = \Sigma l u_1$, $\dot{\sigma}_1 = \Sigma l \dot{u}_1$, $\sigma_2 = \Sigma l u_2$. The conditions for the existence of integrals of linear and angular momentum are satisfied, viz.,

$$\frac{\partial L}{\partial x_1} + \frac{\partial L}{\partial x_2} = 0,$$

and

$$x\frac{\partial L}{\partial y} - y\frac{\partial L}{\partial x} + u_1\frac{\partial L}{\partial v_1} - v_1\frac{\partial L}{\partial u_1} + u_2\frac{\partial L}{\partial v_2} - v_2\frac{\partial L}{\partial u_2} = 0;$$

the terms contributed to angular momentum being

$$x_1\frac{\partial L}{\partial v_1} - y_1\frac{\partial L}{\partial u_1} \text{ and } x_2\frac{\partial L}{\partial v_2} - y_2\frac{\partial L}{\partial u_2}.$$

For the group of odd order the initial term in $2L'_{12}$ is $-\frac{2e_1e_2}{3V^3}\Sigma \ddot{u}_1 u_2$, yielding a force component

$$\xi_2' = \frac{2e_1e_2}{3V^3}\ddot{u}_1 \quad \text{.......................................(11 a).}$$

The sum

$$\Sigma(\xi_1' u_1 + \xi_2' u_2) = \frac{2e_1e_2}{3V^3}\Sigma(\ddot{u}_1 u_2 + u_1\ddot{u}_2) = \frac{d}{dt}\frac{2e_1e_2}{3V^3}(\Sigma\dot{u}_1 u_2 + \Sigma u_1\dot{u}_2) - \frac{4e_1e_2}{3V^3}\Sigma\dot{u}_1\dot{u}_2,$$

and a comparison with

$$\frac{dE'}{dt} + \Sigma(\xi_1' u_1 + \xi_2' u_2) + \text{Radiation} = 0$$

suggests the inference of a term in the expression of the energy $E'$, equal to

$$-\frac{2e_1e_2}{3V^3}(\Sigma\dot{u}_1 u_2 + \Sigma u_1\dot{u}_2) \quad \text{.......................................(11 b)}$$

and a term in radiation equal to $\frac{4e_1e_2}{3V^3}\Sigma\dot{u}_1\dot{u}_2$ .......................................(11 c).

Integration over an infinite sphere (v. *inf.*) confirms the inference which is sufficiently obvious in this simple term. The self-terms which may be inferred from these terms in the mutual action are:

in force $\frac{2e_1^2\ddot{u}_1}{3V^3}$, in radiation $\frac{2e_1^2\Sigma\dot{u}_1^2}{3V^3}$, and in energy $-\frac{2e_1^2}{3V^3}\Sigma u_1\dot{u}_1$ ...........(12).

Taking these in conjunction with the joint terms, we have to this order:

the $x$ component of force on $e_1$ is $\frac{2e_1}{3V^3}(e_1\ddot{u}_1 + e_2\ddot{u}_2)$.......................................(13 a),

the term in radiation of pair is $\frac{2}{3V^3}\Sigma(e_1\dot{u}_1 + e_2\dot{u}_2)^2$ .......................................(13 b),

and the term in expression for energy is

$$-\frac{2}{3V^3}\Sigma(e_1u_1 + e_2u_2)(e_1\dot{u}_1 + e_2\dot{u}_2) \quad \text{.......................................(13 c).}$$

Thus we find that the main term in the odd group gives Larmor's expression for radiation.

The sum of the $x$ components of force on the pair is $2(e_1+e_2)(e_1\ddot{u}_1 + e_2\ddot{u}_2)/3V^3$, which vanishes for a neutral pair, so that only the relative motion is affected. Again, if we treat as approximately true the equations $m_1\dot{u}_1 + m_2\dot{u}_2 = 0$, which are exact when only the dominant term of zero order is taken, the supposition $e_1 : e_2 = m_1 : m_2$ will make the expressions in (13 *a*, *b*, *c*) vanish, i.e. the main terms in radiation and the force associated with it vanish to this order. In effect then we have two types of pair physically differentiated from others, (i) the neutral pair, (ii) the like pair for which $e_1/m_1 = e_2/m_2$.

§ 7. To carry out the differentiations in (5), (6) and (7) it is useful to write $a_1, a_2, a_3, \ldots$ for $(D_1, D_1^2, D_1^3 \ldots) r^2/2$, with $b_1, b_2, b_3, \ldots$ for the operator $D_2$. The opening terms are

$$\left.\begin{array}{l} a_1 = -\Sigma(x_2 - x_1)u_1, \quad a_2 = \Sigma u_1^2 - \Sigma(x_2 - x_1)\dot{u}_1, \quad a_3 = 3\Sigma u_1\dot{u}_1 - \Sigma(x_2 - x_1)\ddot{u}_1, \\ a_4 = 3\Sigma\dot{u}_1^2 + 4\Sigma u_1\ddot{u}_1 - \Sigma(x_2 - x_1)\dddot{u}_1, \quad a_5 = 10\Sigma\dot{u}_1\ddot{u}_1 + 5\Sigma u_1\dddot{u}_1 - \Sigma(x_2 - x_1)u_1^{iv}, \ldots \\ b_1 = \Sigma(x_2 - x_1)u_2, \quad b_2 = \Sigma u_2^2 + \Sigma(x_2 - x_1)\dot{u}_2, \ldots \end{array}\right\} \quad \text{......(14 a)}$$

and we note

$$\left.\begin{array}{l} D_2 a_1 = -\Sigma u_1 u_2, \quad D_2 a_2 = -\Sigma \dot{u}_1 u_2, \quad D_1 b_1 = -\Sigma u_1 u_2, \quad D_1 b_2 = -\Sigma u_1 \dot{u}_2, \ldots \\ \text{making} \qquad r\frac{dr}{dt} = a_1 + b_1, \quad \frac{da_1}{dt} = a_2 - \Sigma u_1 u_2, \quad \frac{da_2}{dt} = a_3 - \Sigma \dot{u}_1 u_2, \ldots \end{array}\right\} \quad \ldots\ldots(14\,b).$$

The numerical coefficients in (14 $a$) are those of the binomial theorem, with the first coefficient halved for letters of even index $a_2, a_4 \ldots$.

In this notation the opening terms of (5) are

$$\frac{2L_{12}}{e_1 e_2} = -\frac{1}{r} - \frac{1}{2V^2 r}\{a_2 - a_1^2/r^2 - 2\Sigma u_1 u_2\} + \frac{1}{3V^3}(a_3 - 3\Sigma \dot{u}_1 u_2)$$

$$-\frac{1}{8V^4 r}\{r^2(a_4 - 4\Sigma \ddot{u}_1 u_2) + 4a_1(a_3 - 2\Sigma \dot{u}_1 u_2) - 4\Sigma u_1 u_2 (a_2 - a_1^2/r^2) + 3(a_2 - a_1^2/r^2)^2\}$$

$$+\frac{1}{30V^5}\{r^2(a_5 - 5\Sigma \dddot{u}_1 u_2) + 10a_1(a_4 - 3\Sigma \ddot{u}_1 u_2) + 10a_3(2a_2 - \Sigma u_1 u_2) - 3a_2 \Sigma \dot{u}_1 u_2\} + \ldots \ldots(15).$$

The opening terms of the symmetrical form (6) are

$$L/e_1 e_2 = -\frac{1}{r} + \frac{\Sigma u_1 u_2 - a_1 b_1/r^2}{2V^2 r}$$

$$+\frac{1}{8V^4 r}\left\{2[(\Sigma u_1 u_2)^2 - a_1^2 b_1^2/r^4] - (a_2 - a_1^2/r^2)(b_2 - b_1^2/r^2) - 2(a_1 \Sigma u_1 \dot{u}_2 + b_1 \Sigma \dot{u}_1 u_2) - 3r^2 \Sigma \dot{u}_1 \dot{u}_2\right\} \ldots \quad \ldots(16),$$

$$\text{or } -\frac{1}{r} + \frac{\Sigma u_1 u_2 + \sigma_1 \sigma_2}{2V^2 r}$$

$$+\frac{1}{8V^4 r}\left\{2[(\Sigma u_1 u_2)^2 - \sigma_1^2 \sigma_2^2] - (\Sigma u_1^2 - \sigma_1^2 - r\dot{\sigma}_1)(\Sigma u_2^2 - \sigma_2^2 + r\dot{\sigma}_2) + 2r(\sigma_1 \Sigma u_1 \dot{u}_2 - \sigma_2 \Sigma \dot{u}_1 u_2) - 3r^2 \Sigma \dot{u}_1 \dot{u}_2\right\} \ldots$$

in the other notation used above.

The opening terms of (7) are

$$\frac{2L'_{12}}{e_1 e_2} = -\frac{2\Sigma \dot{u}_1 u_2}{3V^3} + \frac{1}{15V^5}(a_3 b_2 + 2a_2 \Sigma u_1 \dot{u}_2 + 2r^2 \Sigma \ddot{u}_1 \dot{u}_2 + 7a_1 \Sigma \dot{u}_1 \dot{u}_2 + 3b_1 \Sigma \ddot{u}_1 u_2 - 9\Sigma u_1 u_2 \Sigma \dot{u}_1 u_2) \ldots(17).$$

§ 8. In (15) we have terms of types $a_4$, $a_1 a_3$, $a_2^2$, $a_1^4$ associated with a denominator $V^4$. Thus the expansion (15) will represent a succession of terms of diminishing importance if, $u_1/V$ being of the first order, $r\dot{u}_1/V^2$ is of the second order, $r^2\ddot{u}_1/V^3$ of the third order, and so on. On what does the fulfilment of these conditions turn? If we admit the dominance of the first term $-e_1 e_2/r$ in the K.P. the equations of motion are in the first approximation $m_2 \dot{u}_2 = e_1 e_2 l/r^2 = -m_1 \dot{u}_1, \ldots$. If there is only electromagnetic mass, and $u_1/V$ is not near unity, that mass is a finite multiple of $e_1^2/V^2 c_1$, where $c_1$ is the radius of the electron, and then $r\dot{u}_1/V^2$ will be small if $c_1/r$ is small. If the whole mass is greater than this, the smallness of $r\dot{u}_1/V^2$ can be secured with a less magnitude of orbit. Differentiating the last equation written $e_1 e_2(\dot{x} - 3l\Sigma l\dot{x})/r^3 = -m_1 \ddot{u}_1$, and $r^2\ddot{u}_1/V^3$ is of order $\frac{r\dot{u}_1}{V} \times \frac{\dot{x}}{V}$. Thus when the motion of a pair is considered the successive orders of *acceleration* will conform to the statement above, if the ratio of the *relative* velocity to $V$ is small, and the orbit sufficiently great. In a conical orbit it is generally true that $v^2$ and $fr$ ($v$ and $f$ relative velocity and acceleration) are of the same order, i.e. if $v/V$ is of the first order, $fr/V^2$ is of the second order. But there is a case of exception for a hyperbolic orbit near an apse where $v^2 : fr = e+1$ for an attractive, $e-1$ for a repulsive orbit; and if $e$ is very great the smallness of $fr/V^2$ would not secure $v^2/V^2$ being small. It will appear later that the case of large *absolute* velocities can be dealt with if accelerations are small.

§ 9. We consider now the method of dealing with a kinetic potential in which accelerations of any order appear, and in particular with the two-point case of the conservative system. Let $L$ be a function of $x\,\dot{x}\,\ddot{x}\ldots\overset{(n)}{x}$, and write

$$E = X_1\dot{x} + X_2\ddot{x} + \ldots + X_n\overset{(n)}{x} - L,$$

where $X_1 \ldots X_n$ are to be defined.

Then
$$\frac{dE}{dt} = -\frac{\partial L}{\partial x}\dot{x} - \frac{\partial L}{\partial \dot{x}}\ddot{x} - \ldots - \frac{\partial L}{\partial \overset{(n-1)}{x}}\overset{(n)}{x} - \frac{\partial L}{\partial \overset{(n)}{x}}\overset{(n+1)}{x}$$
$$+ \frac{dX_1}{dt}\dot{x} + \frac{dX_2}{dt}\ddot{x} + \ldots + \frac{dX_n}{dt}\overset{(n)}{x}$$
$$+ X_1\ddot{x} + \ldots + X_{n-1}\overset{(n)}{x} + X_n\overset{(n+1)}{x}.$$

We now define $X_n, X_{n-1}\ldots$ in succession so as to reduce the value of $\dfrac{dE}{dt}$ to the first column. Thus

$$\left.\begin{aligned} &X_n = \frac{\partial L}{\partial \overset{(n)}{x}}, \quad X_{n-1} = \frac{\partial L}{\partial \overset{(n-1)}{x}} - \frac{dX_n}{dt} = \frac{\partial L}{\partial \overset{(n-1)}{x}} - \frac{d}{dt}\frac{\partial L}{\partial \overset{(n)}{x}}, \\ &\text{and if we also introduce} \\ &X_0 = \frac{\partial L}{\partial x} - \frac{dX_1}{dt} = \frac{\partial L}{\partial x} - \frac{d}{dt}\frac{\partial L}{\partial \dot{x}} + \frac{d^2}{dt^2}\frac{\partial L}{\partial \ddot{x}} - \ldots \end{aligned}\right\} \quad\ldots\ldots\ldots\ldots(18),$$

we have $\dfrac{dE}{dt} + X_0\dot{x} = 0$. If there are several variables we write

$$E = \sum_r ({}_rX_1\,\dot{x}_r + {}_rX_2\,\ddot{x}_r + \ldots + {}_rX_n\,\overset{(n)}{x_r}) - L,$$

then
$$\frac{dE}{dt} + \sum_r {}_rX_0\,\dot{x}_r = 0 \quad\ldots\ldots\ldots\ldots(19).$$

Admitting that the quantity $X_0$, occurring in the variation of $L$, represents force in the extended scheme, then $E$, in virtue of satisfying identically the equation (19), is the energy function attached to $L$. If there is force $f_x$ external to the scheme of K.P., then we have $X_0 + f_x = 0$, and $\dfrac{dE}{dt} = \dot{x}f_x$, or more generally $\dfrac{dE}{dt} = \sum_r \dot{x}_r f_{x_r}$; with no such force $E$ is constant.

§ 10. In the two-point system $(x_1y_1z_1)(x_2y_2z_2)$ the total linear momentum $P$ in direction $x$ and the total angular momentum $N$ about the axis of $z$ are given by

$$\left.\begin{aligned} &P = {}_1X_1 + {}_2X_1, \\ \text{and}\quad &N = x_1\,{}_1Y_1 - y_1\,{}_1X_1 + \dot{x}_1\,{}_1Y_2 - \dot{y}_1\,{}_1X_2 + \ddot{x}_1\,{}_1Y_3 - \ddot{y}_1\,{}_1X_3 + \ldots \\ &\quad + x_2\,{}_2Y_1 - y_2\,{}_2X_1 + \dot{x}_2\,{}_2Y_2 - \dot{y}_2\,{}_2X_2 + \ldots \end{aligned}\right\} \quad\ldots\ldots\ldots\ldots(20).$$

We have made use of equivalent K. P.'s differing by $\frac{df}{dt}$; are the quantities $P$ and $N$ given uniquely? Let $f$ be a function of the type of $L$ above, but proceeding only to $\overset{(n-1)}{x_1}, \ldots$; then a K. P.

$$L = \frac{df}{dt} = \dot{x}_1 \frac{\partial f}{\partial x_1} + \ldots + \overset{(n)}{x_1} \frac{\partial f}{\partial \overset{(n-1)}{x_1}} + \dot{y}_1 \frac{\partial f}{\partial y_1} + \ldots + \dot{x}_2 \frac{\partial f}{\partial x_2} + \ldots + \dot{y}_2 \frac{\partial f}{\partial y_2} + \ldots$$

will shew for each $x$ or $y$

$$\left.\begin{aligned} &\frac{\partial L}{\partial \overset{(n)}{x}} = \frac{\partial f}{\partial \overset{(n-1)}{x}}, \quad \frac{\partial L}{\partial \overset{(n-1)}{x}} = \frac{\partial f}{\partial \overset{(n-2)}{x}} + \frac{d}{dt} \frac{\partial f}{\partial \overset{(n-1)}{x}}, \ldots \frac{\partial L}{\partial x} = \frac{d}{dt} \frac{\partial f}{\partial x}; \\ &\text{and therefore} \quad X_n = \frac{\partial f}{\partial \overset{(n-1)}{x}}, \quad X_{n-1} = \frac{\partial f}{\partial \overset{(n-2)}{x}}, \ldots, \quad X_1 = \frac{\partial f}{\partial x} \end{aligned}\right\} \quad \ldots\ldots\ldots\ldots\ldots(21)$$

result from the comparison of columns above. From the last follows $X_0 = 0$, i.e. there is no force, and from $\sum_r (\dot{x}_{r\,r}X_1 + \ddot{x}_{r\,r}X_2 + \ldots) = \frac{df}{dt}$ or $L$ it appears that there is no contribution to energy. The contribution to linear momentum is

$$_1X_1 + {_2X_1} = \frac{\partial f}{\partial x_1} + \frac{\partial f}{\partial x_2},$$

which vanishes if $x_1$ and $x_2$ appear in the combination $(x_2 - x_1)$ as they do in our case. The value of $N$ obtained by substituting in (20) the values of $X_1 \ldots$ given by (21), will also vanish if

$$0 = x_1 \frac{\partial f}{\partial y_1} + \dot{x}_1 \frac{\partial f}{\partial \dot{y}_1} + \ddot{x}_1 \frac{\partial f}{\partial \ddot{y}_1} + \ldots - \left(y_1 \frac{\partial f}{\partial x_1} + \dot{y}_1 \frac{\partial f}{\partial \dot{x}_1} + \ldots\right) + x_2 \frac{\partial f}{\partial y_2} + \ldots - \left(y_2 \frac{\partial f}{\partial x_2} + \ldots\right) \ldots(22),$$

a less simple condition than that for linear momentum, but one satisfied by all the elements used in constructing the formulae for $L$ such as $a_n$, $\Sigma \dot{u}_1 u_2, \ldots$. As an example of the type of proof applicable, suppose the property to hold for $a_3$, and operate with $D_1$ on (22) written with $a_3$ in place of $f$. Now $D_1$ as applied to $a_3$ is

$$\Sigma \left( u_1 \frac{\partial}{\partial x_1} + \dot{u}_1 \frac{\partial}{\partial u_1} + \ddot{u}_1 \frac{\partial}{\partial \dot{u}_1} + \dddot{u}_1 \frac{\partial}{\partial \ddot{u}_1} \right).$$

The difference

$$D_1 \left( x_1 \frac{\partial}{\partial y_1} + u_1 \frac{\partial}{\partial v_1} + \ldots + \ddot{u}_1 \frac{\partial}{\partial \ddot{v}_1} \right) a_3 - \left( x_1 \frac{\partial}{\partial y_1} + \ldots + \ddot{u}_1 \frac{\partial}{\partial \ddot{v}_1} \right) D_1 a_3 = \dddot{u}_1 \frac{\partial a_3}{\partial \ddot{v}_1} = \dddot{u}_1 \frac{\partial a_4}{\partial \dddot{v}_1},$$

or

$$D_1 \left( x_1 \frac{\partial}{\partial y_1} + \ldots \right) a_3 = \left( x_1 \frac{\partial}{\partial y_1} + \ldots + \dddot{u}_1 \frac{\partial}{\partial \dddot{v}_1} \right) a_4,$$

an operator of the same type extended to cover the new form $\dddot{u}$ in $a_4$. The same is true for the section $y_1 \frac{\partial a_3}{\partial x_1} + \ldots$, while for the other two sections $x_2 \frac{\partial}{\partial y_2} \ldots$, $y_2 \frac{\partial}{\partial x_2} \ldots$ the operators are reversible.

In order to make it clear that the value given to $N$ is correct we form

$$\begin{aligned} \frac{dN}{dt} &= x_1 \frac{d\,{_1Y_1}}{dt} + \dot{x}_1 \left( {_1Y_1} + \frac{d}{dt} {_1Y_2} \right) + \ddot{x}_1 \left( {_1Y_2} + \frac{d}{dt} {_1Y_3} \right) + \ldots - \{\ldots\} + \{\ldots\} - \{\ldots\}. \\ &= -x_{1\,1}Y_0 + x_1 \frac{\partial L}{\partial y_1} + \dot{x}_1 \frac{\partial L}{\partial \dot{y}_1} + \ldots - \{\ldots\} + \{\ldots\} - \{\ldots\}, \\ &= -(x_{1\,1}Y_0 - y_{1\,1}X_0) - (x_{2\,2}Y_0 - y_{2\,2}X_0), \end{aligned}$$

in virtue of $L$ satisfying (22). Hence, with ${}_1X_0 + f_{x_1} = 0$, we have

$$\frac{dN}{dt} = x_1 f_{y_1} - y_1 f_{x_1} + x_2 f_{y_2} - y_2 f_{x_2}.$$

Thus the condition (22) plays an important part; as applied to $f$ it ensures K.P.'s which differ by $\frac{df}{dt}$ having the same expression for total angular momentum, while as applied to $L$ it gives the proper dependence of rate of change of angular momentum on external force, i.e. it makes the K.P. system alone conservative in respect to angular momentum.

As an example of derivation, by using (15) or (16) the fourth order term in the value of $E$ is found to be

$$\frac{e_1 e_2}{8V^4 r}\left[2\{(\Sigma u_1 u_2)^2 - a_1^2 b_1^2/r^4\} - (a_2 - a_1^2/r^2)(b_2 - b_1^2/r^2) + (\Sigma u_1 u_2 - 3a_1 b_1/r^2)(a_2 - a_1^2/r^2 + b_2 - b_1^2/r^2) + (a_3 b_1 + a_1 b_3) + (5a_1 - 2b_1)\Sigma \dot{u}_1 u_2 + (5b_1 - 2a_1)\Sigma u_1 \dot{u}_2 + 3r^2 \Sigma(\ddot{u}_1 u_2 + u_1 \ddot{u}_2 - \dot{u}_1 \dot{u}_2)\right] \quad \text{...(22 } a\text{)},$$

while the $x$-component of force on $e_2$ of the same order is

$$\frac{e_1 e_2}{8V^4 r^2}\left[\frac{x}{r}\{3(a_2 - a_1^2/r^2)(a_2 - 5a_1^2/r^2) - 4\Sigma u_1 u_2 (a_2 - 3a_1^2/r^2) + 2a_1(a_3 - 2\Sigma \dot{u}_1 u_2) - r^2(a_4 - 4\Sigma \ddot{u}_1 u_2)\}\right.$$
$$\left. + 4u_1\left\{\frac{b_1}{r}(a_2 - 3a_1^2/r^2) - r\Sigma \dot{u}_1 u_2\right\} - 2r\dot{u}_1\{3(a_2 - a_1^2/r^2) - 2\Sigma u_1 u_2 - 4a_1 b_1/r^2\} - 4r\ddot{u}_1(2a_1 + b_1) - 3r^3 \dddot{u}_1\right] \quad \text{......(22 } b\text{)}.$$

§ 11. I now propose to state what I have found possible in the way of emancipation from conditions as to smallness of velocity in the case of the conservative group of terms, acceleration being still accounted small. In (6) write $x_2 - x_1 = x$ and perform the operation $\nabla^2$ with $xyz$ as space variables. We get

$$\nabla^2 L = -\frac{D_1 D_2}{V^2} L, \text{ or } -\frac{1}{V^2}\frac{d_1}{dt}\frac{d_2}{dt} L, \quad \text{..............................(23)}$$

as a differential equation satisfied by the symmetrical form of K.P. If here we write

$$\frac{d_1}{dt} = \frac{\partial_1}{\partial t} - u_1\frac{\partial}{\partial x} - v_1\frac{\partial}{\partial y} - w_1\frac{\partial}{\partial z}, \qquad \frac{d_2}{dt} = \frac{\partial_2}{\partial t} + u_2\frac{\partial}{\partial x} + v_2\frac{\partial}{\partial y} + w_2\frac{\partial}{\partial z},$$

where $\frac{\partial_1}{\partial t}$ operates on $u_1 \dot{u}_1 \ldots$, and then suppress $\frac{\partial_1}{\partial t}, \frac{\partial_2}{\partial t}$ we obtain

$$\nabla^2 L = \frac{1}{V^2}\left(u_1\frac{\partial}{\partial x} + v_1\frac{\partial}{\partial y} + w_1\frac{\partial}{\partial z}\right)\left(u_2\frac{\partial}{\partial x} + v_2\frac{\partial}{\partial y} + w_2\frac{\partial}{\partial z}\right) L \quad \text{..............(24)}$$

for the quasi-stationary term. A solution is the reciprocal of the square root of a quadric whose coefficients are the minors of the discriminant of $V^2\Sigma x^2 - \Sigma u_1 x \Sigma u_2 x$; this quadric is

$$r^2\{(2V^2 - \Sigma u_1 u_2)^2 - \Sigma u_1^2 \Sigma u_2^2\} + 2(2V^2 - \Sigma u_1 u_2)\Sigma x u_1 \Sigma x u_2 + \Sigma u_1^2 (\Sigma x u_2)^2 + \Sigma u_2^2 (\Sigma x u_1)^2 \text{...(25)},$$

and the term* in the kinetic potential is

$$L = -2e_1 e_2 (V^2 - \Sigma u_1 u_2)/r\sqrt{(2V^2 - \Sigma u_1 u_2 + \sigma_1\sigma_2)^2 - (\Sigma u_1^2 - \sigma_1^2)(\Sigma u_2^2 - \sigma_2^2)} \quad \text{......(26)}.$$

The positive character of the quadric is assured if $1 > \kappa^2 \cos^2\alpha$ where $V^4\kappa^4 = \Sigma u_1^2 \Sigma u_2^2$ and $2\alpha$ is the angle between the directions of the velocities. This is readily seen by

* If in (6) the differentiations are made, and $a_3 \ldots, \Sigma \dot{u}_1 u_2 \ldots$ omitted, the result agrees with (26) when $a_2$, $b_2$ are written for $\Sigma u_1^2$, $\Sigma u_2^2$.

referring the primary quadric to axes two of which are the internal and external bisectors of the directions*.

If the formula (26) is applied to the example of two spheres we get

$$L = -\frac{e_1 e_2}{2V}\int_{-1}^{+1}\frac{(V^2 - n\dot{r}_1\dot{r}_2)\,dn}{\sqrt{V^2(r_1^2 + r_2^2 - 2nr_1r_2) - r_1r_2\dot{r}_1\dot{r}_2(1-n^2)}} = \frac{e_1e_2}{2Vr_1r_2}\left[\sqrt{\phantom{x}}\right]_{-1}^{+1}$$

$$= \frac{e_1e_2}{2Vr_1r_2}\left[\sqrt{V^2(r_2 - r_1)^2} - \sqrt{V^2(r_2+r_1)^2}\right] = -\frac{e_1e_2}{r_2},$$

agreeing with the previous result.

If we seek in (26) a suggestion for the form of the self-term, one course is to make $e_2 = e_1$, $u_2 = u_1$, halve the value and take the mean for all angular positions. This gives

$$\left.\begin{aligned} L_{11} = -\frac{e_1^2(V^2 - \Sigma u_1^2)}{2Vr}\int\frac{1}{\sqrt{V^2 - \Sigma u_1^2 + (\Sigma l u_1)^2}}\frac{d\omega}{4\pi} &= -\frac{e_1^2(V^2 - q_1^2)}{4Vq_1 r}\log\frac{V+q_1}{V-q_1} \\ &= -\frac{3m_1V^2}{8}\frac{V^2 - q_1^2}{Vq_1}\log\frac{V+q_1}{V-q_1}\end{aligned}\right\}\ldots(27),$$

with $q_1^2 = \Sigma u_1^2$ and $m_1V^2 = \dfrac{2e_1^2}{3r}$, $m_1$ being chosen so that the term involving $\Sigma u_1^2$ is $\dfrac{m_1}{2}\Sigma u_1^2$. The inertia in the direction of motion is $\dfrac{d^2L_{11}}{dq_1^2}$, the transverse is $\dfrac{1}{q_1}\dfrac{dL_{11}}{dq_1}$.

If we write $MU = (m_1 + m_2)U = m_1u_1 + m_2u_2$, $u_2 - u_1 = u$ and suppose $u$ small in comparison with $U$, then with

$$L_0 = -\frac{3MV^2}{8}\frac{V^2 - Q^2}{VQ}\log\frac{V+Q}{V-Q}, \qquad Q^2 = \Sigma U^2,$$

we get
$$L_{11} + L_{22} = L_0 + \frac{m_1m_2}{2M^2}\left\{\frac{1}{Q}\frac{dL_0}{dQ}\Sigma u^2 + \left(\frac{d^2L_0}{dQ^2} - \frac{1}{Q}\frac{dL_0}{dQ}\right)\frac{(\Sigma Uu)^2}{Q^2}\right\},$$

as representing terms in $(uvw)$ as far as the second order. Under the same circumstances the main term in (26) is

$$\frac{-e_1e_2(V^2 - \Sigma U^2)}{V\sqrt{(V^2 - \Sigma U^2)r^2 + (\Sigma xU)^2}}.$$

§ 12. We now consider the information given in respect of the non-conservative group by the integrals over an infinite sphere, and at first deal with a single source. The section of a space derivative of potentials which yields finite values is $\dfrac{\partial\psi}{\partial x} = -\dfrac{l_1}{V}\dfrac{\partial\psi}{\partial t}$, where $(x - x_1)/r_1 = l_1$. Consequently at a great distance

$$\left.\begin{aligned} X &= \frac{1}{V}\left(l_1\frac{\partial\psi}{\partial t} - \frac{\partial F}{\partial t}\right), \qquad a = \frac{1}{V}\left(n_1\frac{\partial G}{\partial t} - m_1\frac{\partial H}{\partial t}\right), \\ &\text{and as} \quad \frac{\partial F}{\partial x} + \frac{\partial G}{\partial y} + \frac{\partial H}{\partial z} + \frac{1}{V}\frac{\partial\psi}{\partial t} = 0 \\ &\text{yields} \quad \frac{\partial\psi}{\partial t} = l_1\frac{\partial F}{\partial t} + m_1\frac{\partial G}{\partial t} + n_1\frac{\partial H}{\partial t}\end{aligned}\right\}\ldots\ldots\ldots\ldots(28),$$

* Or by use of velocities 0, 0, $w_1$ and $u_2$, 0, $w_2$, when the condition is $4V^4 - 4V^2w_1w_2 - u_2^2w_1^2 > 0$ or in general terms $(2V^2 - \Sigma u_1u_2)^2 - \Sigma u_1^2\Sigma u_2^2 > 0$.

we have $X = n_1 b - m_1 c, \quad a = m_1 Z - n_1 Y, \quad \Sigma l_1 X = 0, \quad \Sigma l_1 a = 0,$

as for a plane wave. We have also $\Sigma a^2 = \Sigma X^2 = \mathcal{E}$ say, and

$$Yc - Zb = l_1 \mathcal{E}, \quad \Sigma l_1 (Yc - Zb) = \mathcal{E}, \quad l_1 X_x + m_1 X_y + n_1 X_z = l_1 \mathcal{E} = Yc - Zb \quad \ldots(29);$$

viz. since $X_x = \tfrac{1}{2}(Y^2 + Z^2 - X^2) + \tfrac{1}{2}(b^2 + c^2 - a^2)$ or $\mathcal{E} - X^2 - a^2$,

and $X_y = -XY - ab, \quad X_z = -XZ - ac,$

we have $l_1 X_x + m_1 X_y + n_1 X_z = l_1 \mathcal{E} - X \Sigma l_1 X - a \Sigma l_1 a = l_1 \mathcal{E}.$

Thus at the surface of the great sphere a component of Maxwell's stress is also $V \times$ component of momentum-content, or it is $V^{-1} \times$ Poynting's vector.

The sphere is taken with $(x_1 y_1 z_1)$ as centre. If it is supposed fixed, a factor $1 - \sigma_1/V$ is needed for the Doppler effect and the radiation integral is

$$\frac{1}{4\pi} \int \frac{V - \sigma_1}{V} V \mathcal{E} r_1^2 d\omega;$$

if it is treated as moving there is no Doppler effect, and the integral is

$$\frac{1}{4\pi} \int \{ V \Sigma l_1 (Yc - Zb) - v_v \mathcal{E} \} r_1^2 d\omega,$$

which on identifying $v_v$ with $\sigma_1$ again gives

$$\frac{1}{4\pi} \int (V - \sigma_1) \mathcal{E} r_1^2 d\omega.$$

There is the same feature in connexion with the momentum integral which is

$$\frac{1}{4\pi} \int \frac{V - \sigma_1}{V} l_1 \mathcal{E} r_1^2 d\omega;$$

for a moving sphere the form is

$$\frac{1}{4\pi} \int \left\{ \Sigma l X_x - \frac{v_v}{V} (Yc - Zb) \right\} r_1^2 d\omega,$$

reduced to the above by (29).

Now $\psi$ being $e_1 / \{ r_1 - \Sigma (x - x_1) u_1 / V \}$ or $V e_1 / r_1 (V - \sigma_1)$, and $F$ being $u_1 e_1 / r (V - \sigma_1)$, we have at a great distance

$$\frac{\partial \psi}{\partial t_1} = \frac{e_1 V \dot{\sigma}_1}{r_1 (V - \sigma_1)^2}, \qquad \frac{\partial F}{\partial t_1} = \frac{e_1}{r_1} \left\{ \frac{\dot{u}_1}{V - \sigma_1} + \frac{u_1 \dot{\sigma}_1}{(V - \sigma_1)^2} \right\};$$

while
$$\frac{\partial t_1}{\partial t} = \frac{V}{V - \sigma_1};$$

therefore
$$\mathcal{E} = \frac{e_1^2}{r_1^2 (V - \sigma_1)^2} \left\{ \frac{\Sigma \dot{u}_1^2}{(V - \sigma_1)^2} + \frac{2 \dot{\sigma}_1 \Sigma u_1 \dot{u}_1}{(V - \sigma_1)^3} - \frac{(V^2 - \Sigma u_1^2) \dot{\sigma}_1^2}{(V - \sigma_1)^4} \right\};$$

and the radiation integral is

$$\left. \begin{aligned} R &= \frac{e_1^2}{4\pi} \int \left\{ \Sigma \dot{u}_1^2 + \frac{2 \dot{\sigma}_1 \Sigma u_1 \dot{u}_1}{V - \sigma_1} - \frac{(V^2 - \Sigma u_1^2) \dot{\sigma}_1^2}{(V - \sigma_1)^2} \right\} \frac{d\omega}{(V - \sigma_1)^3} \\ &= \frac{2 e_1^2 V}{3 (V^2 - \Sigma u_1^2)^2} \left\{ \Sigma \dot{u}_1^2 + \frac{\Sigma (u_1 \dot{u}_1)^2}{V^2 - \Sigma u_1^2} \right\} \end{aligned} \right\} \quad \ldots\ldots\ldots\ldots\ldots(30).$$

The corresponding integral in the equation of momentum has the extra factor $l_1/V$ in the integrand, and its value is $u_1 R/V^2$.

§ 13. When two sources are in question the joint terms in energy-content are

$$\Sigma (X_1 X_2 + a_1 a_2),$$

and as the difference between $(x-x_1)/r_1$ and $(x-x_2)/r_2$ would involve $r_1^{-1}$ we may take $l_2 = l_1$, and the last relation of (28) makes

$$\Sigma a_1 a_2 = \Sigma X_1 X_2 - \Sigma l_1 X_1 \Sigma l_1 X_2 = \Sigma X_1 X_2,$$

so that the terms in energy-content are $2\Sigma X_1 X_2$. The value of $\Sigma X_2 X_2$ is

$$\frac{e_1 e_2}{r_1^2 (V-\sigma_1)^2 (V-\sigma_2)^2} \left\{ \Sigma \dot{u}_1 \dot{u}_2 + \frac{\dot{\sigma}_1 \Sigma u_1 \dot{u}_2}{V-\sigma_1} + \frac{\dot{\sigma}_2 \Sigma \dot{u}_1 u_2}{V-\sigma_2} - \frac{(V^2 - \Sigma u_1 u_2)\dot{\sigma}_1 \dot{\sigma}_2}{(V-\sigma_1)(V-\sigma_2)} \right\} \quad \text{.........(31).}$$

The Doppler effect, or effect of moving surface of integration, will be dealt with as follows. For the flux of momentum localized at $e_1$, and for a corresponding flux of energy, integration is taken over a sphere with centre at $e_1$, with the Doppler factor for that centre and the integrand $\Sigma X_1 X_2$. As $X_2$ depends on $u_2$ a function of $t_2$, the phase-difference between $t_2$ and $t_1$ will be calculated for points on the sphere, and thereby the expression will be brought to one time. Corresponding parts due to a sphere with $e_2$ as centre will give the flux of momentum localized at $e_2$, and a corresponding flux of energy. These fluxes each using half the energy-content will constitute the total. It is proposed to give evidence that this method ensures the adaptation to the forces stated to be essential.

The times are connected by

$$V(t_2 - t_1) = r_1 - r_2 = \Sigma (x_2 - x_1)(x - x_1)/r_1 = \Sigma l_1 (x_2 - x_1),$$

in which $x_2$ is $x_2(t_2)$. If now we use $x_2, u_2, \ldots$ for $x_2(t_1), \dot{x}_2(t_1), \ldots$,

$$x_2(t_2) = x_2 + (t_2 - t_1) u_2 + \frac{(t_2-t_1)^2}{2!} \dot{u}_2 + \ldots$$

and

$$V(t_2 - t_1) = \Sigma l_1 \left\{ x_2 - x_1 + u_2 (t_2 - t_1) + \frac{\dot{u}_2}{2!}(t_2-t_1)^2 + \ldots \right\};$$

or, with $x$ now for relative coordinate $x_2 - x_1$, and $\sigma_0 = \Sigma l_1 x$,

$$(t_2 - t_1)\left\{ V - \sigma_2 - \frac{\dot{\sigma}_2}{2!}(t_2-t_1) - \frac{\ddot{\sigma}_2}{3!}(t_2-t_1)^2 \ldots \right\} = \sigma_0,$$

i.e.

$$t_2 - t_1 = \frac{\sigma_0}{V-\sigma_2} + \frac{\sigma_0^2 \dot{\sigma}_2}{2(V-\sigma_2)^3} + \frac{\sigma_0^3}{2(V-\sigma_2)^4}\left\{ \frac{\ddot{\sigma}_2}{3} + \frac{\dot{\sigma}_2^2}{V-\sigma_2} + \frac{\sigma_0 \ddot{\sigma}_2}{12(V-\sigma_2)} \right\} + \ldots \quad \text{......(32).}$$

With this difference we have

$$u_2(t_2) = u_2 + \frac{\dot{u}_2 \sigma_0}{V-\sigma_2} + \frac{\ddot{u}_2 \sigma_0^2}{2(V-\sigma_2)^2} + \frac{\dddot{u}_2 \sigma_0^3}{6(V-\sigma_2)^3} + \frac{\dot{u}_2 \sigma_0^2 \dot{\sigma}_2}{2(V-\sigma_2)^3} + \ldots$$

and similar forms, and in particular

$$V - \sigma_2(t_2) = V - \sigma_2 - \frac{\sigma_0 \dot{\sigma}_2}{V-\sigma_2}$$

gives

$$\frac{1}{V-\sigma_2(t_2)} = \frac{1}{V-\sigma_2}\left\{ 1 + \frac{\sigma_0 \dot{\sigma}_2}{(V-\sigma_2)^2} + \ldots \right\}.$$

§ 14. Thus denoting by $R(e_1)$ this first section of the joint term in radiation, we have on introducing the Doppler factor

$$
\begin{aligned}
R(e_1) = \frac{e_1 e_2}{4\pi} \int \frac{d\omega}{(V-\sigma_1)(V-\sigma_2)^2} \Bigg[ & \left\{ \Sigma \dot{u}_1 \dot{u}_2 + \frac{\dot{\sigma}_1 \Sigma u_1 \dot{u}_2}{V-\sigma_1} + \frac{\dot{\sigma}_2 \Sigma \dot{u}_1 u_2}{V-\sigma_2} - \frac{(V^2 - \Sigma u_1 u_2)\dot{\sigma}_1 \dot{\sigma}_2}{(V-\sigma_1)(V-\sigma_2)} \right\} \\
& + \frac{\sigma_0}{V-\sigma_2} \left\{ \Sigma \dot{u}_1 \ddot{u}_2 + \frac{\dot{\sigma}_1 \Sigma u_1 \ddot{u}_2}{V-\sigma_1} + \frac{\ddot{\sigma}_2 \Sigma \dot{u}_1 u_2 + \dot{\sigma}_2 \Sigma \dot{u}_1 \dot{u}_2}{V-\sigma_2} + \frac{\Sigma u_1 \dot{u}_2 \dot{\sigma}_1 \dot{\sigma}_2 - (V^2 - \Sigma u_1 u_2)\dot{\sigma}_1 \ddot{\sigma}_2}{(V-\sigma_1)(V-\sigma_2)} \right\} \\
& + \frac{\sigma_0^2}{2(V-\sigma_2)^2} \{ \Sigma \dot{u}_1 \dddot{u}_2 + \dots \} \\
& + \frac{2\sigma_0 \dot{\sigma}_2}{(V-\sigma_2)^2} \left\{ \Sigma \dot{u}_1 \dot{u}_2 + \frac{\dot{\sigma}_1 \Sigma u_1 \dot{u}_2}{V-\sigma_1} \right\} + \frac{3\sigma_0 \dot{\sigma}_2}{(V-\sigma_2)^2} \left\{ \frac{\dot{\sigma}_2 \Sigma u_1 \dot{u}_2}{V-\sigma_2} - \frac{(V^2 - \Sigma u_1 u_2)\dot{\sigma}_1 \dot{\sigma}_2}{(V-\sigma_1)(V-\sigma_2)} \right\} + \dots \Bigg] \quad \dots (33),
\end{aligned}
$$

in which the lines after the first are due to phase-differences. The $x$-component of flux of momentum is the same integral with the extra factor $l_1/V$. For the moment we require these evaluated to the order $V^{-5}$ for use in conjunction with force on $e_1$ and work done on $e_1$. The evaluation is

$$
F_x(e_1) = \frac{e_1 e_2}{15 V^5} [4x \Sigma \dot{u}_1 \ddot{u}_2 + (3u_1 + 7u_2) \Sigma \dot{u}_1 \dot{u}_2 - \dot{u}_1 \{b_3 - 3\Sigma u_1 \dot{u}_2\} + 2\dot{u}_2 \Sigma \dot{u}_1 (u_2 - u_1) - \ddot{u}_2 \Sigma x \dot{u}_1] \dots (34a),
$$

$$
\begin{aligned}
R(e_1) = \frac{2e_1 e_2}{3V^3} \Sigma \dot{u}_1 \dot{u}_2 + \frac{e_1 e_2}{15 V^5} [ & (2\Sigma u_1^2 + 9\Sigma u_1 u_2 + 9b_2) \Sigma \dot{u}_1 \dot{u}_2 + 2r^2 \Sigma \dot{u}_1 \dddot{u}_2 + (11b_1 - 3a_1) \Sigma \dot{u}_1 \ddot{u}_2 \\
& + (a_2 - \Sigma u_1^2)(b_4 - 3\Sigma u_1 \ddot{u}_2) + \Sigma \dot{u}_1 u_2 (b_3 + 9\Sigma u_1 \dot{u}_2) - 2\Sigma u_1 \dot{u}_1 (b_3 - 2\Sigma u_1 \dot{u}_2)] \dots (34b).
\end{aligned}
$$

Now, working from the formula $2L_{21}$ in (17),

$$
\begin{aligned}
\xi_1' = \frac{2e_1 e_2}{3V^3} \ddot{u}_2 + \frac{e_1 e_2}{15 V^5} [ & -x(b_5 - 5\Sigma u_1 \dddot{u}_2) + 10 u_2 \Sigma u_1 \ddot{u}_2 + 10 \dot{u}_2 b_3 \\
& + 10 \ddot{u}_2 (2b_2 - \Sigma u_1 u_2) + 5\dot{u}_2 (3b_1 + a_1) + 2r^2 \ddddot{u}_2] \dots (35a),
\end{aligned}
$$

and

$$
\begin{aligned}
\Sigma \xi_1' u_1 = \frac{2e_1 e_2}{3V^3} \Sigma u_1 \ddot{u}_2 + \frac{e_1 e_2}{15 V^5} [ & a_1 (b_5 - 5\Sigma u_{12}) + 10 b_3 \Sigma u_1 \dot{u}_2 + 20 b_2 \Sigma u_1 \ddot{u}_2 \\
& + 5(3b_1 + a_1) \Sigma u_1 \dddot{u}_2 + 2r^2 \Sigma u_1 \ddddot{u}_2] \dots (35b).
\end{aligned}
$$

Thus the sum

$$
\begin{aligned}
F_x(e_1) + \xi_1' = \frac{d}{dt} \Bigg[ \frac{2e_1 e_2}{3V^3} \dot{u}_2 + \frac{e_1 e_2}{15 V^5} \{ & -x(b_4 - 4\Sigma u_1 \ddot{u}_2) - u_1 (b_3 - 3\Sigma u_1 \dot{u}_2) + u_2 (b_3 + 7\Sigma u_1 \dot{u}_2) \\
& + \dot{u}_2 (9b_2 + 2\Sigma u_1 u_2 - \Sigma u_1^2) + \ddot{u}_2 (11b_1 + a_1) + 2r^2 \dddot{u}_2 \} \Bigg] \dots (36) \\
= -\frac{dp_1'}{dt} \text{ say}.
\end{aligned}
$$

and the sum

$$
\begin{aligned}
R(e_1) + \Sigma \xi_1' u_1 = \frac{d}{dt} \Bigg[ \frac{2e_1 e_2}{3V^3} \Sigma u_1 \dot{u}_2 + \frac{e_1 e_2}{15 V^5} \{ & 2r^2 \Sigma u_1 \dddot{u}_2 + (11b_1 + a_1) \Sigma u_1 \ddot{u}_2 + a_1 (b_4 - 4\Sigma u_1 \ddot{u}_2) \\
& + b_3 \Sigma u_1 (u_2 - u_1) + \Sigma u_1 \dot{u}_2 (2\Sigma u_1^2 + 9\Sigma u_1 u_2 + 9b_2) \} \Bigg] \dots (37) \\
= -\frac{dE'(e_1)}{dt} \text{ say}.
\end{aligned}
$$

Next we observe that with these values of $E'$ and $p_1'$ we have the relation $E'(e_1) = \Sigma p_1' u_1$. We now form the quantity $\Sigma u_1 F_x(e_1)$ and subtract it from $R(e_1)$, i.e. we take from the total rate of loss of energy that which is of mechanical order, due to flux of momentum, and denote by $S(e_1)$ the pure radiation which is left. We find

$$
\begin{aligned}
\Sigma u_1 F_x(e_1) = \frac{e_1 e_2}{15 V^5} [ & -4a_1 \Sigma \dot{u}_1 \ddot{u}_2 + (3\Sigma u_1^2 + 7\Sigma u_1 u_2) \Sigma \dot{u}_1 \dot{u}_2 - \Sigma u_1 \dot{u}_1 (b_3 - \Sigma u_1 \dot{u}_2) \\
& + 2\Sigma u_1 \dot{u}_2 \Sigma \dot{u}_1 u_2 - \Sigma u_1 \ddot{u}_2 \Sigma x \dot{u}_1],
\end{aligned}
$$

and thence

$$S(e_1) = \frac{2e_1e_2}{3V^3}\Sigma\dot{u}_1\dot{u}_2 + \frac{e_1e_2}{15V^5}[-\Sigma x\dot{u}_1(b_4 - 4\Sigma u_1\ddot{u}_2) + 2r^2\Sigma\dot{u}_1\ddot{u}_2 + (11b_1 + a_1)\Sigma\dot{u}_1\ddot{u}_2$$
$$+ \Sigma\dot{u}_1\dot{u}_2(9b_2 + 2\Sigma u_1u_2 - \Sigma u_1^2) - \Sigma u_1\dot{u}_1(b_3 - 3\Sigma u_1\dot{u}_2) + \Sigma\dot{u}_1u_2(b_3 + 7\Sigma u_1\dot{u}_2)] \ldots (38).$$

$S$ is linear in respect to $(\dot{u}_1\dot{v}_1\dot{w}_1)$ and we find that $S(e_1) = -\Sigma p_1'\dot{u}_1$.

We have then a scheme of relations

$$\left.\begin{aligned} \xi_1' + F_x(e_1) + \frac{dp_1'}{dt} = 0, \quad \Sigma\xi_1'u_1 + R(e_1) + \frac{dE'(e_1)}{dt} = 0 \\ E'(e_1) = \Sigma p_1'u_1, \quad S(e_1) = -\Sigma p_1'\dot{u}_1, \quad R(e_1) = S(e_1) + \Sigma u_1F_x(e_1) \end{aligned}\right\} \ldots\ldots\ldots\ldots(39).$$

Thus the conditions laid down that the sum of force and flux of momentum, and the sum of rate of working of force and rate of radiation, should give time-rates of momentum and energy, are satisfied to the fifth order by our method of treating the integrals over infinite spheres. The position is confirmed by the simple relations which are then found to obtain between energy and momenta, and also between pure radiation and momenta. I have also evaluated the flux of momentum to the seventh order, where the flux contains upwards of 50 terms (in the compressed notation with $a$, $b$), and this condition is again satisfied. The radiation condition of this order was not examined, but I feel little doubt that the whole scheme of (39) is exact.

But if this scheme is of general validity, it is evidently possible to proceed in a different order, viz. to find integrals $R$ and $F$, infer $S$ and thence $p'$ and so $\xi'$ and $E'$, that is to construct the whole scheme of forces of non-conservative order with expressions for momentum and energy, from the integrals at infinity. The advantage of this is that these integrals can be evaluated without reference to the magnitude of $u_1/V$, $u_2/V$; and there remains only the condition that the phase-differences should be small enough to admit of treatment by expansion.

The application of this method to the self-terms gives the result

$$\left.\begin{aligned} S &= \frac{2e_1^2}{3V(V^2 - \Sigma u_1^2)}\left\{\Sigma\dot{u}_1^2 + \frac{(\Sigma u_1\dot{u}_1)^2}{V^2 - \Sigma u_1^2}\right\}, \\ p' &= -\frac{2e_1^2}{3V(V^2 - \Sigma u_1^2)}\left\{\dot{u}_1 + \frac{u_1\Sigma u_1\dot{u}_1}{V^2 - \Sigma u_1^2}\right\}, \\ E' &= -\frac{2e_1^2V\Sigma u_1\dot{u}_1}{3(V^2 - \Sigma u_1^2)^2}, \\ \xi' &= \frac{2e_1^2}{3V(V^2 - \Sigma u_1^2)}\left[\ddot{u}_1 + \frac{u_1\Sigma u_1\ddot{u}_1 + 3\dot{u}_1\Sigma u_1\dot{u}_1}{V^2 - \Sigma u_1^2} + \frac{3u_1(\Sigma u_1\dot{u}_1)^2}{(V^2 - \Sigma u_1^2)^2}\right] \end{aligned}\right\} \ldots\ldots\ldots\ldots\ldots(40).$$

with $R$ and $F$ as given above in (30). This agrees with the value of $\xi'$ given by Abraham.

§ 15. The integration for the joint terms can be carried out by exact methods, and we propose to give this integration for the main terms, i.e. omitting those dependent on phase-differences. The integrals in

$$R(e_1) = \frac{e_1e_2}{4\pi}\int\frac{d\omega}{(V - \sigma_1)(V - \sigma_2)^2}\left\{\Sigma\dot{u}_1\dot{u}_2 + \frac{\dot{\sigma}_1\Sigma u_1\dot{u}_2}{V - \sigma_1} + \frac{\dot{\sigma}_2\Sigma\dot{u}_1u_2}{V - \sigma_2} - \frac{(V^2 - \Sigma u_1u_2)\dot{\sigma}_1\dot{\sigma}_2}{(V - \sigma_1)(V - \sigma_2)}\right\}$$

are derivable by differentiation with respect to $u_1$ or $u_2$ from a fundamental integral* over a sphere, viz.

$$f = \frac{1}{4\pi}\int \frac{(V^2 - \Sigma u_1 u_2)\, d\omega}{(V - \sigma_1)(V - \sigma_2)} = \frac{C}{2\sqrt{C^2 - AB}} \log \frac{C + \sqrt{C^2 - AB}}{C - \sqrt{C^2 - AB}} = \frac{1}{2x} \log \frac{1+x}{1-x} \quad \text{......(41)}$$

with the notation

$$A = V^2 - \Sigma u_1^2, \quad B = V^2 - \Sigma u_2^2, \quad C = V^2 - \Sigma u_1 u_2, \quad x^2 = (C^2 - AB)/C^2$$

In differentiating it is also convenient to use $y^2 = AB/C^2 = 1 - x^2$, and then from

$$\frac{1}{y}\frac{\partial y}{\partial u_1} = -\frac{u_1}{A} + \frac{u_2}{C}, \qquad \frac{1}{y}\frac{\partial y}{\partial u_2} = -\frac{u_2}{B} + \frac{u_1}{C},$$

we form

$$\frac{\partial f}{\partial u_1} = f_1\left(-\frac{u_1}{A} + \frac{u_2}{C}\right), \quad \text{where } f_1 = y\frac{df}{dy}.$$

Extending the notation, i.e. putting $f_2 = y\dfrac{df_1}{dy}$, ..., we have

$$1 + f_1 = y^2(f + f_1), \quad f_2 = y^2(2f + 3f_1 + f_2), \quad f_3 = y^2(4f + 8f_1 + 5f_2 + f_3), \quad \dots,$$

leading to $\quad 2 + 2f_1 - f_2 = -y^2(f_1 + f_2)$ and $2f_2 - f_3 = -y^2(2f_1 + 3f_2 + f_3)$,

which with $y^2 = AB/C^2$ can be used to modify forms of the result†.

The integration yields:

$$R(e_1) = \frac{e_1 e_2 V}{2}\left[\frac{\Sigma \dot{u}_1 \dot{u}_2}{C}\left(\frac{f_2 - f_1}{B} - \frac{f_2 + f_1}{C}\right) - \frac{\Sigma u_1 \dot{u}_1 \, \Sigma u_2 \dot{u}_2}{AB}\left(\frac{f_2 + f_3}{C} + \frac{2f_2 - f_3}{B}\right)\right.$$

$$+ \frac{\Sigma \dot{u}_1 u_2 \, \Sigma u_1 \dot{u}_2}{BC}\left(\frac{2f_2 - f_3}{A} + \frac{f_3 - f_1}{C}\right) + \frac{\Sigma u_1 \dot{u}_1 \, \Sigma u_1 \dot{u}_2}{AC}\left(\frac{f_2 + f_3}{C} + \frac{f_2 - f_3}{B}\right)$$

$$\left. + \frac{\Sigma u_2 \dot{u}_2 \, \Sigma \dot{u}_1 u_2}{BC}\left(\frac{f_1 + 2f_2 + f_3}{C} + \frac{2f_2 - f_3}{B}\right)\right] \quad \text{..............................(42 a),}$$

* Using velocities $(0, 0, w_1)$, $(u_2, 0, w_2)$ and putting $V = 1$, we require

$$\frac{1}{4\pi}\int_{-1}^{+1}\int_0^{2\pi} \frac{dn\, d\phi}{(1 - w_1 n)(1 - w_2 n - u_2\sqrt{1 - n^2}\cos\phi)} \quad \text{or} \quad \frac{1}{2}\int_{-1}^{+1} \frac{dn}{(1 - w_1 n)\sqrt{1 - u_2^2 - 2nw_2 + n^2(u_2^2 + w_2^2)}},$$

for which we quote

$$\int \frac{dn}{(\gamma - n)\sqrt{a - 2bn + cn^2}} = \frac{1}{\sqrt{a - 2b\gamma + c\gamma^2}} \log \frac{a - bn + \gamma(cn - b) + \sqrt{(a - 2b\gamma + c\gamma^2)(a - 2bn + cn^2)}}{\gamma - n};$$

and the integral is then

$$\frac{1}{2\sqrt{aw_1^2 - 2bw_1 + c}} \log \frac{(1 + w_1)\{(aw_1 - b) + (c - bw_1) + (1 - w_2)\}\sqrt{aw_1^2 - 2bw_1 + c}}{(1 - w_1)\{(aw_1 - b) - (c - bw_1) + (1 + w_2)\}\sqrt{aw_1^2 - 2bw_1 + c}}$$

$$= \frac{1}{2\sqrt{aw_1^2 - 2bw_1 + c}} \log \frac{1 - w_1 w_2 + \sqrt{aw_1^2 - 2bw_1 + c}}{1 - w_1 w_2 - \sqrt{aw_1^2 - 2bw_1 + c}}.$$

To get the last form cross-multiply the fractions and use

$$aw_1 - b + w_1(c - bw_1) = (w_1 - w_2)(1 - w_1 w_2).$$

Then note that $\quad aw_1^2 - 2bw_1 + c = (1 - w_1 w_2)^2 - (1 - w_1^2)(1 - u_2^2 - w_2^2)$,

write $V$ where 1 appears, and the general character of the result is evident.

† Opening terms of expansions:

$$f = 1 + \frac{x^2}{3} + \frac{x^4}{5} + \frac{x^6}{7} \dots, \qquad f_1 = -\frac{2}{3} - \frac{2x^2}{15} - \frac{2x^4}{35} \dots, \qquad f_2 = \frac{4}{15} - \frac{4x^2}{105} - \frac{4x^4}{105} \dots, \qquad f_3 = \frac{8}{105} + \frac{8x^2}{105} + \frac{8x^4}{385} \dots$$

$$F_x(e_1) = \frac{e_1 e_2}{2V}\left[u_1\left\{-\frac{\Sigma \dot{u}_1 \dot{u}_2}{C^2}(f_1 + f_2) + \frac{\Sigma \dot{u}_1 u_2 \, \Sigma u_1 \dot{u}_2}{ABC}(2f_2 - f_3) - \frac{\Sigma u_1 \dot{u}_1 \, \Sigma u_2 \dot{u}_2}{ABC}(f_2 + f_3)\right.\right.$$

$$+ \frac{\Sigma u_1 \dot{u}_1 \, \Sigma u_1 \dot{u}_2}{AC^2}(f_2 + f_3) + \left.\frac{\Sigma \dot{u}_1 u_2 \, \Sigma u_2 \dot{u}_2}{BC^2}(f_1 + 2f_2 + f_3)\right\}$$

$$+ u_2\left\{\frac{\Sigma \dot{u}_1 \dot{u}_2}{BC}(f_2 - f_1) + \frac{\Sigma \dot{u}_1 u_2 \, \Sigma u_1 \dot{u}_2}{BC^2}(f_2 - f_1) - \frac{\Sigma u_1 \dot{u}_1 \, \Sigma u_2 \dot{u}_2}{AB^2}(2f_2 - f_3)\right.$$

$$+ \frac{\Sigma u_1 \dot{u}_1 \, \Sigma u_1 \dot{u}_2}{ABC}(f_2 - f_3) + \left.\frac{\Sigma \dot{u}_1 u_2 \, \Sigma u_2 \dot{u}_2}{B^2C}(2f_2 - f_3)\right\}$$

$$\left. + \frac{\dot{u}_1}{C}\left(\frac{\Sigma u_2 \dot{u}_2}{B} - \frac{\Sigma u_1 \dot{u}_2}{C}\right)(f_1 + f_2) + \frac{\dot{u}_2}{B}\left(\frac{\Sigma \dot{u}_1 u_2}{C} - \frac{\Sigma u_1 \dot{u}_1}{A}\right)f_2\right] \quad \ldots\ldots\ldots\ldots(42\,b);$$

thence

$$S(e_1) = \frac{e_1 e_2}{2V}\left[\frac{\Sigma \dot{u}_1 \dot{u}_2}{B}(2 + f_1) + \frac{\Sigma u_1 \dot{u}_1 \, \Sigma u_2 \dot{u}_2}{BC}(f_1 + f_2) + \frac{\Sigma \dot{u}_1 u_2 \, \Sigma u_1 \dot{u}_2}{BC}(f_2 - f_1)\right.$$

$$\left. - \frac{\Sigma u_1 \dot{u}_1 \, \Sigma u_1 \dot{u}_2}{C^2}(f_1 + f_2) + \frac{\Sigma \dot{u}_1 u_2 \, \Sigma u_2 \dot{u}_2}{B^2}(2 + 2f_1 - f_2)\right] \quad \ldots\ldots(42\,c),$$

and

$$E'(e_1) = -\frac{e_1 e_2 V}{2}\left[(2 + 2f_1 - f_2)\frac{\Sigma u_2 \dot{u}_2}{B^2} + \frac{\Sigma u_2 \dot{u}_2}{BC}(f_1 + f_2) + \frac{\Sigma u_1 \dot{u}_2}{BC}(f_2 - f_1) - \frac{\Sigma u_1 \dot{u}_2}{C^2}(f_1 + f_2)\right]$$

$$\ldots\ldots(42\,d).$$

If here we write $F_x(e_1) = \alpha u_1 + \beta u_2 + \alpha_1 \dot{u}_1 + \beta_1 \dot{u}_2$, then

$$R(e_1) - S(e_1) = \Sigma u_1 F_x(e_1) = V^2(\alpha + \beta) - (A\alpha + C\beta - \alpha_1 \Sigma u_1 \dot{u}_1 - \beta_1 \Sigma u_1 \dot{u}_2),$$

if for $\Sigma u_1^2$ and $\Sigma u_1 u_2$ we introduce $V^2 - A$, $V^2 - C$ to compare with the forms in (42). It will be found that $R(e_1) = \dot{V}^2(\alpha + \beta)$. Thus we can go a step further, and say that the evaluation of $F$ is the only integration needed. This does not appear to be affected when the phase terms are introduced, though the form of $F$ is thereby extended to

$$F_x(e_1) = \gamma x + (\alpha u_1 + \alpha_1 \dot{u}_1) + (\beta u_2 + \beta_1 \dot{u}_2 + \beta_2 \ddot{u}_2 + \ldots).$$

An example of this is seen if we look back to the fifth order value of $F_x(e_1)$, when we find that $V^2(\alpha + \beta)$ gives correctly the third order term in $R$. Also in obtaining the value of $F_x(e_1)$ to the seventh order I found that these terms give correctly the value of the fifth order term in $R$, a further addition to the evidence.

§ 16. A case in which the formulae are much simplified is got by writing

$$U = (m_1 u_1 + m_2 u_2)/(m_1 + m_2), \quad u = u_2 - u_1$$

and then assuming that $u$ is negligible in comparison with $U$, while $\dot{U}$ is negligible in comparison with $\dot{u}$. Thus in effect we put $U$ for $u_1$ or $u_2$, and $\dot{U} = 0$ makes $\dot{u}_1 = -m_2 \dot{u}/M$, $\dot{u}_2 = m_1 \dot{u}/M$, where $M = m_1 + m_2$. As a first approximation we then get

$$\left.\begin{aligned} S(e_1) &= -\frac{2e_1 e_2 m_1 m_2}{3M^2 V}\left\{\frac{\Sigma \dot{u}^2}{V^2 - \Sigma U^2} + \frac{(\Sigma U \dot{u})^2}{(V^2 - \Sigma U^2)^2}\right\} \\ R(e_1) &= -\frac{2e_1 e_2 m_1 m_2 V}{3M^2}\left\{\frac{\Sigma \dot{u}}{(V^2 - \Sigma U^2)^2} + \frac{(\Sigma U \dot{u})^2}{(V^2 - \Sigma U^2)^3}\right\} \\ F_x(e_1) &= -\frac{2e_1 e_2 m_1 m_2 U}{3M^2 V}\left\{\frac{\Sigma \dot{u}^2}{(V^2 - \Sigma U^2)^2} + \frac{(\Sigma U \dot{u})^2}{(V^2 - \Sigma U^2)^3}\right\} \end{aligned}\right\} \quad \ldots\ldots\ldots\ldots\ldots\ldots(43).$$

with the same values for $S(e_2)$, $R(e_2)$ and $F_x(e_2)$. Hence with the self-terms the totals of radiation, and the sum $F_x$ concerned in total linear momentum, are

$$R = \frac{2(e_1 m_2 - e_2 m_1)^2 V}{3M^2}\left\{\frac{\Sigma \dot{u}^2}{(V^2 - \Sigma U^2)^2} + \frac{(\Sigma U \dot{u})^2}{(V^2 - \Sigma U^2)^3}\right\}, \quad S = \frac{R(V^2 - \Sigma U^2)}{V^2}, \quad F_x = \frac{UR}{V^2} \ \ldots(44).$$

These quantities vanish for a like pair with the property $e_1 : e_2 = m_1 : m_2$; while for a neutral pair $(e_1 m_2 - e_2 m_1)^2$, $M^2 = e_1^2$ or $e_2^2$.

The sum of the self and joint terms in the momentum of $e_1$ is

$$p'(e_1) = \frac{2e_1(e_1 m_2 - e_2 m_1)}{3MV(V^2 - \Sigma U^2)}\left\{\dot{u} + \frac{U \Sigma U \dot{u}}{V^2 - \Sigma U^2}\right\} \ldots\ldots\ldots\ldots(45\,a),$$

and therefore the total linear momentum is

$$\frac{2(e_1 + e_2)(e_1 m_2 - e_2 m_1)}{3MV(V^2 - \Sigma U^2)}\left\{\dot{u} + \frac{U \Sigma U \dot{u}}{V^2 - \Sigma U^2}\right\} \ldots\ldots\ldots\ldots(45\,b),$$

which vanishes in both special cases.

§ 17. It is proposed to examine in detail how the primary motion is modified by the main dissipative terms, those of the third order, which we shall treat as small quantities while gnoring the conservative terms of the second order. The equations of motion are

$$m_2 \dot{u}_2 = \frac{e_1 e_2 l}{r^2} + \frac{2e_2}{3V^3}(e_1 \ddot{u}_1 + e_2 \ddot{u}_2), \qquad m_1 \dot{u}_1 = -\frac{e_1 e_2 l}{r^2} + \frac{2e_1}{3V^3}(e_1 \ddot{u}_1 + e_2 \ddot{u}_2) \ \ldots\ldots(46),$$

and the integral of linear momentum is

$$m_1 u_1 + m_2 u_2 - \frac{2}{3V^3}(e_1 + e_2)(e_1 \dot{u}_1 + e_2 \dot{u}_2) = P \ldots\ldots\ldots\ldots(47\,a).$$

For a neutral pair this reduces to the primary form, and the relative motion only is affected. For a like pair with the property $e_1 : e_2 = m_1 : m_2$, $(47\,a)$ yields

$$m_1 u_1 + m_2 u_2 - P = \{(m_1 u_1 + m_2 u_2)_{t=0} - P\}\, e^{2(e_1 + e_2)^2 t / 3MV^3},$$

which requires $m_1 u_1 + m_2 u_2 = P$ initially and therefore always, unless we are to suppose that linear momentum can increase indefinitely. In this case the problem is reduced to that given by the primary terms.

But if we take the problem of (46) with general values of $e_1$ and $e_2$, and introduce in the small terms of $(47\,a)$ values resulting from the primary solution, we get

$$m_1 u_1 + m_2 u_2 - \frac{2e_1 e_2 (e_1 + e_2)}{3V^3}\left(\frac{e_2}{m_2} - \frac{e_1}{m_1}\right)\frac{l}{r^2} = P \ \ldots\ldots\ldots\ldots(47\,b),$$

and for relative acceleration

$$\dot{u} = \frac{e_1 e_2 l}{r^2}\left(\frac{1}{m_1} + \frac{1}{m_2}\right) + \frac{2m_1 m_2 \ddot{u}}{3MV^3}\left(\frac{e_2}{m_2} - \frac{e_1}{m_1}\right)^2.$$

With the notation

$$p = -\frac{2e_1 e_2}{3V^3}\left(\frac{e_2}{m_2} - \frac{e_1}{m_1}\right)^2$$

we have

$$l\dot{u} + m\dot{v} + n\dot{w} = \frac{e_1 e_2}{r^2}\left(\frac{1}{m_1} + \frac{1}{m_2}\right) + \frac{2p\dot{r}}{r^4} \ldots\ldots\ldots\ldots(48),$$

and

$$x\dot{v} - y\dot{u} = -\frac{p}{r^3}(xv - yu), \ \ldots$$

shewing central acceleration and areal change in convenient forms.

With multipliers $u_2$, $u_1$ ... in (46) the equation of energy is

$$\frac{d}{dt}\left[\frac{m_1}{2}\Sigma u_1^2+\frac{m_2}{2}\Sigma u_2^2+\frac{e_1e_2}{r}-\frac{2}{3V^3}\Sigma(e_1u_1+e_2u_2)(e_1\dot{u}_1+e_2\dot{u}_2)\right]=-\frac{2}{3V^3}\Sigma(e_1\dot{u}_1+e_2\dot{u}_2)^2 \quad\text{...(49 a)}.$$

Write $e_1\dot{u}_1+e_2\dot{u}_2=\left(\frac{e_2}{m_2}-\frac{e_1}{m_1}\right)\frac{e_1e_2l}{r^2}$ in the small terms of $E$, and introduce the momenta from (47) and we get a simplified expression for $E$ in the form

$$E=\frac{\Sigma P^2}{2M}+\frac{m_1m_2}{2M}\Sigma u^2+\frac{e_1e_2}{r}+\frac{m_1m_2p\dot{r}}{Mr^2} \quad\text{.........................(49 b)}.$$

The rate of radiation $\frac{2}{3V^3}\Sigma(e_1\dot{u}_1+e_2\dot{u}_2)^2$ has the value $-\frac{pe_1e_2}{r^4}$ or in fact

$$\frac{dE}{dt}=\frac{pe_1e_2}{r^4} \quad\text{.............................................(50)}.$$

The loss of energy by radiation in one revolution is

$$-pe_1e_2\int_0^T\frac{dt}{r^4}=-\frac{pe_1e_2}{k}\int_0^{2\pi}\frac{d\theta}{r^2}=-\frac{pe_1e_2}{kl^2}\int_0^{2\pi}(1+e\cos\theta)^2\,d\theta=-\frac{e_1e_2p\pi}{kl^2}(2+e^2)\quad\text{...(51 a)},$$

and the mean rate of loss, on division by the period, is

$$-\frac{pe_1e_2(2+e^2)}{2a^4(1-e^2)^{\frac{5}{2}}} \quad\text{.......................................(51 b)},$$

where in the small terms we have assumed the results of an elliptic orbit, $k$ being the constant of areal description, and $e$, $a$, and $l$ having the usual meanings in Conics.

§ 18. So far it has been possible to dispense with an account of the actual deformation of orbit: this we proceed to consider in the simpler case where the motion is plane, and the position in polar coordinates is defined by

$$\frac{dh}{d\theta}=-\frac{p}{r},\qquad \frac{1}{r}+\frac{d^2}{d\theta^2}\frac{1}{r}=-\frac{e_1e_2}{h^2}\left(\frac{1}{m_1}+\frac{1}{m_2}\right)+\frac{3p}{hr}\frac{d}{d\theta}\frac{1}{r} \quad\text{.................(52)}.$$

The solution based on the fundamental solution $h=k$, $\frac{1}{l}=-\frac{e_1e_2}{k^2}\left(\frac{1}{m_1}+\frac{1}{m_2}\right)$ $\frac{l}{r}=1+e\cos\theta$, when radiation is neglected, is

$$\left.\begin{aligned}
&h=k(1+\kappa)-\frac{p}{l}(\theta+e\sin\theta)\\
&\frac{l}{r}+\frac{d^2}{d\theta^2}\frac{l}{r}=1-2\kappa+\frac{p}{kl}(2\theta-e\sin\theta-3e^2\sin\theta\cos\theta)\\
&\frac{l}{r}=1+e\cos\theta-2\kappa+\alpha\cos\theta+\beta\sin\theta+\frac{p}{kl}\left(2\theta+\frac{e\theta}{2}\cos\theta+e^2\sin\theta\cos\theta\right)
\end{aligned}\right\}\text{...(53 a)}.$$

The orbit can be made to agree with $l=r(1+e\cos\theta)$ in respect of $r$, $\frac{dr}{d\theta}$, $\frac{d^2r}{d\theta^2}$, and to have $h=k$ at $\theta=0$, by taking $\alpha=\kappa=0$, $\beta=-\frac{p}{kl}\left(2+\frac{e}{2}+e^2\right)$. It appeared however in dealing with the period and some mean values, that simpler formulae are obtained by making this agreement correspond to $\theta=\pi$ in the middle of the circuit. We make use of this determination, which gives

$$\kappa=\frac{p\pi}{kl},\qquad \alpha=-\frac{ep\pi}{2kl},\qquad \beta=\frac{p}{kl}\left(2-\frac{e}{2}+e^2\right) \quad\text{...................(53 b)}.$$

In (53 $a$) we distinguish between a cyclic and non-cyclic perturbation, i.e. we write $\frac{l}{r} = 1 + e\cos\theta + U_0$ in the first revolution, and

$$\frac{l}{r} = 1 + e\cos\theta + U_0 + U_c \quad \text{or} \quad 1 + e\cos\theta + U_0 + \frac{p\pi}{kl}(4 + e\cos\theta)$$

in the second revolution.

Thus we suppose that the $\theta$ which occurs in $U_0$ is always between 0 and $2\pi$.

If then the constants of the modified ellipse in the second revolution are $l'$, $k'$, $e'$,

$$\left.\begin{aligned} &\frac{1}{l'} = \frac{1}{l}\left(1 + \frac{4p\pi}{kl}\right), \quad \frac{e'}{l'} = \frac{e}{l}\left(1 + \frac{p\pi}{kl}\right), \quad e' = e\left(1 - \frac{3p\pi}{kl}\right) \\ &a' = a\left\{1 - \frac{2p\pi(2+e^2)}{kl(1-e^2)}\right\}, \quad k' = k\left(1 - \frac{2p\pi}{kl}\right) \end{aligned}\right\} \quad \text{.............. (54).}$$

The period in the orbit is

$$T = \int_0^{2\pi} \frac{r^2 d\theta}{h} = \frac{l^2}{k}\int_0^{2\pi}\left[1 + \frac{3p(\pi - \theta)}{kl(1 + e\cos\theta)} + \dots\right]\frac{d\theta}{(1 + e\cos\theta)^2},$$

where terms containing $\sin\theta$ are not written. This gives a normal value to the period, viz.

$$2\pi l^2/k(1-e^2)^{\frac{3}{2}} = 2\pi a^{\frac{3}{2}}/\sqrt{-e_1e_2M/m_1m_2}.$$

The term to be added for the next revolution is

$$\frac{l^2}{k}\int_0^{2\pi}\frac{(-6p\pi)\,d\theta}{kl(1+e\cos\theta)^3} \quad \text{or} \quad -\frac{6p\pi^2 l(2+e^2)}{k^2(1-e^2)^{\frac{5}{2}}},$$

i.e. the next period is

$$\frac{2\pi l^2}{k(1-e^2)^{\frac{3}{2}}}\left\{1 - \frac{3p\pi(2+e^2)}{kl(1-e^2)}\right\} = 2\pi a'^{\frac{3}{2}}\sqrt{-e_1e_2M/m_1m_2},$$

which conforms to the modification of the fundamental ellipse.

In the same way the mean potential energy taken for a revolution is

$$\frac{e_1e_2}{T}\int\frac{dt}{r} = \frac{e_1e_2}{T}\int\frac{r\,d\theta}{h} = \frac{e_1e_2l}{kT}\int_0^{2\pi}\left[1 - \frac{p(\pi-\theta)}{2kl} + \frac{3p(\pi-\theta)}{2kl(1+e\cos\theta)} - \dots\right]\frac{d\theta}{1+e\cos\theta}$$

$$= \frac{2\pi e_1e_2l}{kT\sqrt{1-e^2}} = \frac{e_1e_2}{a} \quad \text{.................................................... (55).}$$

The mean kinetic energy of relative motion or $\frac{m_1m_2}{2MT}\int\Sigma u^2dt$ is

$$\frac{m_1m_2k}{2MT}\int_0^{2\pi}\left[2 - \frac{1-e^2}{1+e\cos\theta} - \frac{p(\pi-\theta)}{kl}\left\{1 - \frac{5+e^2}{1+e\cos\theta} + \frac{3(1-e^2)}{(1+e\cos\theta)^2}\right\} + \dots\right]\frac{d\theta}{1+e\cos\theta}$$

$$= \frac{m_1m_2k}{2MT\sqrt{1-e^2}} = \frac{m_1m_2k^2}{2Mal} = -\frac{e_1e_2}{2a} \quad \text{.............................. (56).}$$

These values are normal, and it is readily shewn that in the next revolution the means retain the same form with the altered values of $a$. These values give as the loss of energy in a period, $\frac{e_1e_2}{2}\left(\frac{1}{a} - \frac{1}{a'}\right) = -\frac{e_1e_2p\pi(2+e^2)}{kl^2}$ in agreement with (51 $a$).

§ 19. The loss of energy is associated with a contraction of the major axis, and a diminution of eccentricity. The mean kinetic energy of relative motion is increased by an amount equal to the loss of energy by radiation, and the mean potential energy bears the double loss. This is an immediate consequence of the maintenance of the relations

$$2T_r + U_r = 0 \quad \text{and} \quad T_r + U_r = E_r \quad \text{..............................(57)},$$

where $T_r$ and $U_r$ are mean values of relative kinetic energy and of potential energy.

A brief statement may be made in respect to hyperbolic orbits, where the total radiation may be calculated, a quantity which in the theory of point-charges represents the radiation due to a collision. For the attractive case make the comparison with the fundamental orbit $l = r(e\cos\theta + 1)$ at $\theta = 0$, i.e. take $\alpha = 0 = \kappa$ and $\beta = \frac{p}{kl}\left(2 + \frac{e}{2} + e^2\right)$. The radiation in the complete orbit is

$$-\frac{2pe_1e_2}{kl^2}\int_0^{\pi-\theta_0}(1 + e\cos\theta)^2\,d\theta = -\frac{pe_1e_2}{kl^2}\{(2 + e^2)(\pi - \theta_0) + 3e\sin\theta_0\} \quad \text{......(58 a)},$$

where $\theta_0 < \frac{\pi}{2}$ and $e\cos\theta_0 = 1$.

For the repulsive case we get

$$-\frac{pe_1e_2}{kl^2}\{(2 + e^2)\,\theta_0 - 3e\sin\theta_0\} \quad \text{..............................(58 b)}.$$

The value of the constant $p$ when $e_1 = -e_2 = \epsilon$ is $p = \frac{2\epsilon^4}{3V^3}\frac{M^2}{m_1^2m_2^2}$, and if we suppose that $e_2$ is a negative electron, so that the ratio $m_2 : m_1$ is small, then $p = \frac{2\epsilon^2}{3V^3}\left(\frac{\epsilon}{m_2}\right)^2$. The coefficient of the bracket in (58 a) is then $\frac{2\epsilon^6}{3m_2^2kl^2V^3}$, or, if $l$ is eliminated, it is

$$\frac{2}{3m_2^4V^3}\left(\frac{\epsilon^2}{k}\right)^5.$$

For the elliptic orbit formula, (51 b) is

$$\frac{2 + e^2}{3V^3(1 - e^2)^{\frac{3}{2}}}\left(\frac{\epsilon}{m_2}\right)^2\left(\frac{\epsilon^2}{a}\right)^4,$$

and the number $\frac{p}{kl}$ occurring in the orbital changes is $\frac{2}{3}\left(\frac{\epsilon^2}{m_2kV}\right)^3$.

It may be noticed that if $a$ is taken inversely proportional to temperature $\theta$, then the kinetic energy of relative motion is proportional to $\theta$, and the rate of radiation to $\theta^4$. This is no doubt a significant point; but application to the thermodynamics of radiation probably demands a statistical treatment of a large number of elements and the groups which they can form.

§ 20. I have also solved the problem of the primary motion as modified by the terms of second order, which when $e$ is finite gives more trouble in the integrations. The results are of the type found in discussing the question of a modification of gravity as applied to explain secular changes in the orbit. For the present purpose their importance seems hardly commensurate with the space needed to prove them.

In conclusion a brief review of some points in the paper is added.

The fact that a form $2L_{12}$ is used to give force on $e_2$ while $2L_{21}$ gives force on $e_1$, raises a presumption that the kinetic potential has not a normal character. The departure from normal type is not easy to locate exactly, without the use of an expansion proceeding by powers of $V^{-1}$, an expansion certainly valid for a wide range of motions. It is then definitely located in the section of terms of odd order, and these terms only are concerned in radiation.

The even groups in $L_{12}$ and $L_{21}$ are shewn to have dynamical equivalence (§ 4), and to form an entirely conservative system if treated alone (§ 10). This conservative section, when acceleration is negligible, admits of a quasi-stationary kinetic potential without assumption as to smallness of velocities.

Electromagnetic force is known in respect to odd or even sections: in the conservative section an expression for energy follows at once, in the dissipative section not until radiation is evaluated. Closely connected with this is the question of localizing momentum and energy, i.e. distinguishing the parts attached to the two charges, a problem solved for the conservative section by the use of the kinetic potential. For the dissipative section it is necessary to call in the aid of fluxes at infinity. In view of the fact that two centres are concerned it is not immediately evident how this flux is to be treated. But the fact that we are using information furnished by two methods implies that a correspondence is to be found which will be a criterion of correct treatment of the flux. The agreement of two methods of reduction to a one-time system is involved.

This adjustment is in fact attained as far as the approximation extends, and it is presumed that the scheme of relations (39) so deduced has general validity. This carries with it the possibility of presenting the radiation from two sources in a form free from limitation as to the magnitude of velocities; and also of deducing expressions for the terms of odd order in energy, momentum, and force directly from the fluxes at infinity. The integrals concerned are all derivable from one fundamental integral (41) involving the sources in a symmetrical way.

It will be noted that the argument in general deals with joint or product terms in the action of two point-charges. The transition to self-terms for the dissipative section presents no difficulties: in the conservative section infinite values would appear. It is only in this connexion that the necessity of giving finite though small dimensions to the electron arises. The method used in the text does not postulate definite structure, but I think the decision in the matter must be left to experimental evidence as to the ratio of two inertias in the case of rapid motion.

(*Received* 1 December 1915; *read* 21 February 1916.)

# XII. *The Field and the Cordon of a Plane Set of Points.*
## *An Essay in Proving the Obvious.*

BY ERIC H. NEVILLE.

CONTENTS.



### 1. *Introduction and summary.*

The objects of this paper do not include the introduction to its readers of the sets of points with which principally the paper is concerned, for with these sets every mathematician is well acquainted: the set which is here called the field of a set $\Gamma$ is the least convex set containing $\Gamma$, and I have given the name of cordon to the boundary of this set.

In a number of simple cases the field of a set $\Gamma$ may be defined in terms of centres of mass: thus if $\Gamma$ is a curve of finite length, and line density is given to $\Gamma$, the position of the centre of mass of $\Gamma$ depends on the distribution of density, but there is a certain region $F`\Gamma$ such that however the density may be arranged, continuously or discontinuously, provided only that the density is nowhere negative, the centre of mass lies in $F`\Gamma$, and such moreover that whatever point $x$ of $F`\Gamma$ is chosen some arrangement of density can be contrived which brings the centre of mass to $x$; it is this region $F`\Gamma$ which is in this case the field of $\Gamma$. Again, if $\Gamma$ consists of a finite number of points in a plane, there is only one polygon, in the Euclidean sense of the word, which has all its vertices at points of $\Gamma$, has no reentrant angles, and has all the points of $\Gamma$ in its interior or upon its sides; in this case the sides of this polygon compose the cordon, and the field consists of the cordon together with its interior. From these two examples the importance both of the field and of the cordon will be recognised; the field is involved in almost all mean-value theorems in multiple integration, and the cordon, apart from its relation to the field, is to be found in use in the Newton parallelogram for dealing with branches of a curve and in the Puiseux diagram in connection with linear differential equations.

The bulk of this paper is concerned solely with plane sets of points, and what is offered is a definition of the field of a set $\Gamma$ in a form at once applicable to plane sets of all kinds, limited and unlimited, open and closed, and convenient for the development of the properties of a field. The essence of the definition consists in the use of a geometrical element of which, as far as I know, the introduction is a novelty; this element, which I call a leaf, consists of a point together with all the points of some line through it which lie on one side of it and all the points of the plane which lie on one side of this line; no set of a quite simple character can reasonably be called a half-plane, and there is no set simpler than a leaf which more closely resembles its complement. A point $x$ is said to be outside the field of $\Gamma$ if there is a leaf which includes $x$ and includes no point belonging to $\Gamma$, the field is the complement of the set composed of points outside the field, and the cordon is the common boundary of the field and its complement. The cordon in general lies partly inside and partly outside the field, and an account is given of properties distinguishing the two portions of the cordon.

After such explanation as seems necessary of the notation adopted, the paper proceeds to exact definitions of the sets of special kinds which are used, it being stated in each case what part, if any, of the boundary is included; the theorems of which use is to be made are enunciated, proofs being omitted, but in two cases where existence theorems (11·37, 11·42) are required constructions are given. With the definitions of the primary (12·11) and secondary (13·12) chords and of the cross points (14·11) of a set we come nearer to our main topic, and the digression to the definition of a convex set (15·11) is not irrelevant. Two ideas of which much use is subsequently made are next explained; roughly, the excluding angle (16·12) of a set for a point is the angle of the biggest sectors which have the point for vertex and have no points of the set within them, and a sector $\Sigma$ is called a limiting excluding sector (16·13) for a set $\Gamma$ if $\Sigma$ itself contains no points of $\Gamma$ but every sector with the same vertex as $\Sigma$ of which $\Sigma$ is a proper part does contain at least one point of Γ: in this connection it must be noticed that in anticipation of this stage our formal definitions (9·11, 9·12) of a sector are such that no existent sector contains any part of its boundary. Each point of a plane falls with respect to a set Γ into one of three classes, denoted by $U$‘Γ, $V$‘Γ, and $W$‘Γ, according as the excluding angle of Γ for it is greater than, equal to, or less than, $\pi$, and the only propositions in the paper of which the proofs are tiresome are those (18·35, 19·45) which describe characteristics of the sets $V$‘Γ and $W$‘Γ. These propositions established we define the field (20·21, 21·11) and the cordon (21·14) of a set, and we have no difficulty in demonstrating so many properties of these sets as to render it evident both that our definitions are well adapted for the development of formal proofs, and that the subject is one in which what is obvious is true. Something is said of the nature of the field of a set with respect to any relation in space of any kind, and of the properties of the field in Euclidean space of any finite number of dimensions, and the paper concludes with suggesting a line of research.

The principal theorems regarding plane sets of which the paper contains proofs may be summarised as follows:

I. The field of a plane set Γ is composed of the points of Γ and the points lying on primary and secondary chords of Γ; if the set Γ is connected or is the sum of two connected parts every point of the field belongs to Γ or to a primary chord of Γ: (26·15, 26·31, 12·14).

II. The field of a plane set Γ is itself a convex set containing Γ, and is the set formed of all the points common to all convex sets containing Γ: (22·34, 21·22, 26·59).

III. If a set is enlarged by the addition of any part of its boundary, the cordon is unaltered: (24·23).

IV. The points inside the cordon of Γ are the cross points of Γ and the points lying on secondary chords of Γ, and they are the points for which the excluding angle is less than $\pi$: (23·36, 19·45).

V. A point lies outside the cordon of Γ if there is a straight line parting it from Γ: (23·59).

VI. If the cordon consists of the whole of one straight line or of two parallel straight lines, the excluding angle for every point outside the cordon is equal to $\pi$, but in all other cases in which the cordon exists the excluding angle for every point outside the cordon is greater than $\pi$: (23·65).

VII. A point of the cordon of Γ which belongs to the field of Γ belongs either to Γ or to a primary chord of Γ: (24·14, 24·16).

VIII. A point of the cordon of Γ for which Γ has an excluding angle greater than $\pi$ either belongs to Γ or is a limiting point of Γ: (24·14, 24·15).

IX. Unless Γ consists of only one point, every point of the cordon of Γ is a limiting point of points on primary chords of Γ: (24·13).

## 2. *Explanation of logical symbols used in this paper and not described in Chapter I of the Introduction to "Principia Mathematica."*

Except in the use of the letters $B$, $C$, $D$, $F$ and in the absence of any sign of assertion, the logical notation of this paper is that of *Principia Mathematica*. There are only a few symbols used here which are not among those explained in Chapter I of the Introduction to that work, and by describing the use of these I hope to render the paper intelligible to anyone familiar with that Chapter.

If a symbol is introduced to give brevity to a few proofs without any implication that the idea associated with it has permanent value, the definition introducing the symbol is called a temporary definition, and the letters Df which distinguish a permanent definition are replaced by Dft followed by an indication in square brackets of the extent to which the new symbol is to be used.

If $\kappa$ is a class whose members are classes, that class whose members are all the terms which belong to every member of $\kappa$ is called the product of $\kappa$ and denoted by $p`\kappa$, and that class is called the sum of $\kappa$ and denoted by $s`\kappa$ which is such that a term belongs to $s`\kappa$ if and only if there is at least one member of $\kappa$ to which it belongs:

2·21 $$p`\kappa = \hat{x}\{(\alpha) . \alpha \epsilon \kappa \supset x \epsilon \alpha\} \quad \text{Df,}$$

·22 $$s`\kappa = \hat{x}\{(\exists \alpha) : \alpha \epsilon \kappa . x \epsilon \alpha\} \quad \text{Df.}$$

The number of members which a class $\alpha$ contains is denoted by $\text{Nc}`\alpha$; thus if $\Gamma$ is a set of points, $\text{Nc}`\Gamma$ denotes the number of points in the set, while if $\kappa$ is a class of sets $\text{Nc}`\kappa$ is the number of sets belonging to $\kappa$, but the number of points concerned in the constitution of $\kappa$ is $\text{Nc}`s`\kappa$.

The authors of *Principia Mathematica* have occasion to use two different pieces of symbolism for the one idea of the class formed of those terms which have a given relation $R$ to a single term; in our applications, the terms in question are in fact always sets of points. If $R$ is a relation which is not in all cases one-many which holds between one set of points and another, the class of sets whose members are all the sets which have the relation $R$ to $\Gamma$ can be denoted by either $\overrightarrow{R}`\Gamma$ or $(\text{sg}`R)`\Gamma$. The relations which we denote by single letters are all as a matter of fact one-many relations: if $R$ denotes one of these relations, there is only one set which has the relation $R$ to $\Gamma$, this set can be denoted by $R`\Gamma$, and the class of which this set is the only member is $\iota`R`\Gamma$. Thus almost all of the cases in which we require a symbol for a class of referents are cases in which the use of an arrow is inconvenient or its appearance unsightly: for example, to print an arrow above the group of letters exlf, which occurs in one of our most important definitions, would not only be inelegant in formulae but also affect the spacing of lines if the resulting combination occurred in the body of the text. To limit the groups of letters used in expressing relations to groups with which we do not object to printing an arrow would in many cases prevent such a choice of letters as assists the memory, and uniformly we adopt the alternative notation; for example, the group of symbols that occurs in 20·22 is $(\text{sg}`\text{exlf})`\Gamma$, and in 4·3 we use $(\text{gs}`\epsilon)`x$ rather than $\overleftarrow{\epsilon}`x$ for the class formed of the sets to which $x$ belongs.

There is one feature of the notation of *Principia Mathematica* to which attention is not called in the first Chapter, although it is recognised later as a natural outcome of the theory of classes there explained. Possession of any property which objects of any kind, individuals, classes, relations, may have, can be treated symbolically as membership of the class of objects possessing that property. For example, we have presently to define what is meant by the assertion that a set of points is united, and the symbol which is introduced to correspond to the idea of a united set is a symbol not for an adjective but for the whole class of united sets: we write $\Gamma \epsilon \mathrm{Ud}$ just as we write $x \epsilon \alpha$, but while in the one case we read "$x$ is an $\alpha$", in the other we may read simply "$\Gamma$ is united"; the distinction between possession of a property and membership of a class is one of language alone, and needs no embodiment in logical symbols. An immediate consequence of this result is that, if we wish to denote that a set has one or other of a number of properties or has several properties simultaneously, we can use the ordinary notation for the logical sum or for the logical product of classes: for example, the numbers 0 and 1 are themselves regarded as classes, and the condition $\Gamma \epsilon 0 \cup 1$ is equivalent to the condition $\Gamma \epsilon 0 \,.\, \mathsf{v} \,.\, \Gamma \epsilon 1$; similarly important properties of sets are expressed by the terms complete set and congregate, Cp is used for the class of complete sets and Cg for the class of congregates, and to write $\Gamma \epsilon \mathrm{Cp} \cap \mathrm{Cg}$ is to assert that $\Gamma$ is a complete congregate.

It is chiefly as a form of shorthand that the notation of *Principia Mathematica* is required in this paper. Nevertheless to frame definitions in the form which this notation is best adapted to express is the surest guarantee that the ideas involved are logically precise, and in this connection I owe thanks to Prof. Whitehead himself for criticism of my manuscript which has led to considerable modification in the formal definitions contained in the earlier sections.

### 3. *Definitions adopted from the general theory of sets of points.*

The explanations yet given are virtually extracts from *Principia Mathematica*, accounts of general logical symbolism. Next must be described the notation used to express certain ideas peculiar to the theory of sets of points but common to all parts of this theory, and this is done quite briefly, the reader being referred for a fuller discussion of the ideas involved to a paper shortly to appear in the *Acta Mathematica*. Throughout I use $t, u, v, w, x, y, z$ for individual points, $xy$ for the distance between $x$ and $y$, $\Gamma, \Delta, \Theta, \Phi, \Psi$ for sets of any kind, V for the universe of points, that is, except in sections 27—29, for the set composed of all the points of the plane in which our sets are supposed to lie. $\Lambda$ denotes the null set of points: that is to say, to write $\Gamma = \Lambda$ is to assert that there are no points satisfying the conditions that define membership of $\Gamma$, so that for example the formula $\Gamma \cap \Delta = \Lambda$ expresses that $\Gamma$ and $\Delta$ have no common point; $\exists ! \Gamma$ denotes that $\Gamma$ is not null, and is the contradictory of $\Gamma = \Lambda$.

If $\Gamma$ is any set, I denote by $C`\Gamma$ the complement of $\Gamma$, the set formed of all points which do not belong to $\Gamma$, by $D`\Gamma$ the derivative of $\Gamma$, the set formed of all the limiting points of $\Gamma$, and by $G`\Gamma$ the set obtained by completing $\Gamma$, that is, the set $\Gamma \cup D`\Gamma$ obtained by adding to $\Gamma$ all those of its limiting points which do not belong to it; also I denote by $Y`\Gamma$ the edge of $\Gamma$, that is, the set $\Gamma \cap D`C`\Gamma$ formed of the members of $\Gamma$ which are limiting points of the complement $C`\Gamma$, and by $B`\Gamma$ the set known as the boundary of $\Gamma$, that is, the sum of the edges of $\Gamma$ and its complement, and I describe the set $\Gamma - Y`\Gamma$, which is the same as $\Gamma - D'C'\Gamma$, as the set obtained by clipping $\Gamma$, and denote this set by $H'\Gamma$:

3·21 $$C'\Gamma = V - \Gamma \quad \text{Df},$$

·22 $$D'\Gamma = \hat{x}\{\rho > 0 \supset_\rho (\exists y) . y \in \Gamma . 0 < xy < \rho\} \quad \text{Df},$$

·23 $$G'\Gamma = \Gamma \cup D'\Gamma \quad \text{Df},$$

·24 $$Y'\Gamma = \Gamma \cap D'C'\Gamma \quad \text{Df},$$

·25 $$B'\Gamma = Y'\Gamma \cup Y'C'\Gamma \quad \text{Df},$$

·26 $$H'\Gamma = \Gamma - D'C'\Gamma \quad \text{Df}.$$

A set is said to be dense if every one of its points is a limiting point, complete if it contains all its limiting points, a domain if it has no edge, and limited* if it does not extend to infinity; the contractions used are shewn in the following formal definitions:

·31 $$\text{Ds} = \hat{\Gamma}\{\Gamma \subset D'\Gamma\} \quad \text{Df},$$

·32 $$\text{Cp} = \hat{\Gamma}\{D'\Gamma \subset \Gamma\} \quad \text{Df},$$

·33 $$\text{Dom} = \hat{\Gamma}\{Y'\Gamma = \Lambda\} \quad \text{Df},$$

·34 $$\text{Lm} = \hat{\Gamma}\{(\exists \rho) . y, z \in \Gamma \supset_{y,z} yz < \rho\} \quad \text{Df}.$$

From ·33 and ·24

·35 $$\Gamma \in \text{Dom} . \equiv . D'C'\Gamma \subset C'\Gamma,$$

and so from ·32

·36 $$\Gamma \in \text{Dom} . \equiv . C'\Gamma \in \text{Cp},$$

a property that might be used to define one of the two classes Dom, Cp in terms of the other:

·37 $$\text{Dom} = C''\text{Cp}.$$

The definitions

·38 $$\text{Cl} = \text{Cp} \cap \text{Lm} \quad \text{Df},$$

·39 $$\text{Pf} = \text{Ds} \cap \text{Cp} \quad \text{Df},$$

shew the useful combinations of properties associated with the words closed and perfect.

Formally the null set belongs to all of the classes defined in the last paragraph; sometimes it is convenient to express briefly that a set is an *existent* set with the property characteristic of a class of sets in which the null set is included, and to this end we add ex to the symbol of the class† to denote that the null set has been removed: thus we write

·41 $$\text{Domex} = \text{Dom} - \iota'\Lambda \quad \text{Df},$$

and so on, and we have

·42 $$\Gamma \in \text{Domex} \equiv . \exists ! \Gamma . \Gamma \in \text{Dom}.$$

I propose to say that a set $\Gamma$ is a congregate§ if in every expression of $\Gamma$ as the sum of two sets of which neither is null, the sets obtained by completing these sets have at

* It is only in certain kinds of space that this property defines a limited set, but Euclidean space is of one such kind, whatever the number of dimensions.

† Compare *Principia Mathematica*, *60·02.

§ The reader will observe that a set that is connex in Cantor's sense is not necessarily a congregate; a pair of conjugate hyperbolas is a connex set formed of four distinct congregates.

least one point in common, and to describe a set Γ as united if every pair of points contained in Γ belongs to some closed congregate contained in Γ, writing

$$\cdot 51 \qquad \mathrm{Cg} = \hat{\Gamma} \{\Gamma = \Theta \cup \Phi \,.\, \exists\,!\,\Theta \,.\, \exists\,!\,\Phi : \supset_{\Theta,\Phi} \exists\,!\, G`\Theta \cap G`\Phi\} \quad \mathrm{Df},$$

$$\cdot 52 \qquad \mathrm{Ud} = \hat{\Gamma} \{y, z \in \Gamma \supset_{y,z} : (\exists\Delta) \,.\, \Delta \in \mathrm{Cl} \cap \mathrm{Cg} \,.\, y, z \in \Delta \,.\, \Delta \subset \Gamma\} \quad \mathrm{Df}.$$

Any existent set Γ can be expressed as the sum of a class of mutually exclusive united sets, and these sets I call the cells of Γ. To keep the definitions as simple as possible, it is best to write

$$\cdot 53 \qquad K_\Gamma`x = \hat{y} \{(\exists\Delta) \,.\, \Delta \in \mathrm{Cl} \cap \mathrm{Cg} \,.\, x, y \in \Delta \,.\, \Delta \subset \Gamma\} \quad \mathrm{Df},$$

without the hypothesis that $x$ belongs to Γ: this definition gives

$$\cdot 54 \qquad x \in C`\Gamma \supset K_\Gamma`x = \Lambda,$$

but to define the class of cells of Γ in such a way as not to include the null set we have only to take the definition

$$\cdot 55 \qquad \kappa`\Gamma = K_\Gamma``\Gamma \quad \mathrm{Df},$$

which is an abbreviated form of

$$\cdot 56 \qquad \kappa`\Gamma = \hat{\mathrm{K}} \{(\exists x) \,.\, x \in \Gamma \,.\, \mathrm{K} = K_\Gamma`x\}.$$

The number of cells of Γ, that is, the number Nc‘κ‘Γ, is precisely the number which common sense assigns to the distinct parts of which Γ is composed; the null set has no cells. The notation of ·53 is essential to the elegance of the definition ·55, but is inadequate when the set whose cells are under consideration is given not directly but by a construction of any sort, and we therefore write also

$$\cdot 57 \qquad K`(x, \Gamma) = K_\Gamma`x \quad \mathrm{Df}.$$

There is one idea which in its general form most naturally depends on united sets or on cells, which we use in a particular case. A set Θ is said to part two sets Γ, Δ if Γ and Δ are both contained in the complement of Θ, but no cell of this complement includes members of both Γ and Δ; an equivalent definition is that Θ parts Γ and Δ if Γ and Δ are both contained in the complement of Θ and if every closed congregate which includes members of both Γ and Δ includes a member of Θ. In the second form the definition is formally independent of the definition of a united set. Taking the first definition,

$$\cdot 61 \qquad C`\Phi \text{ part } (\Gamma, \Delta) = :. \Gamma \cup \Delta \subset \Phi : y \in \Gamma \,.\, z \in \Delta \,.\, \supset_{y,z} K_\Phi`y \cap K_\Phi`z = \Lambda \quad \mathrm{Df},$$

the equivalence is expressed in the theorem

$$\cdot 62 \qquad \Theta \text{ part } (\Gamma, \Delta) \equiv :. \Theta \cap (\Gamma \cup \Delta) = \Lambda : \Psi \in \mathrm{Cl} \cap \mathrm{Cg} \,.\, \exists\,!\,\Psi \cap \Gamma \,.\, \exists\,!\,\Psi \cap \Delta \,.\, \supset_\Psi \exists\,!\,\Psi \cap \Theta.$$

We have in this paper to consider the relation of a point $x$ to a set Γ when there is a straight line $h$ which does not pass through $x$ and is such that no point of Γ lies on $h$ or on the same side of $h$ as $x$, and we adapt the general notation to this case, expressing the relation by $h$ part $(\iota`x, \Gamma)$; but we can easily give a definition applicable only to special cases, putting

$$\cdot 63 \qquad h \in \mathrm{Stl} \supset : h \text{ part } (\iota`x, \Gamma) = . \, y \in \Gamma \supset_y \exists\,!\, h \cap x - y \quad \mathrm{Df},$$

where Stl stands for straight line and $x - y$ for the set of points lying between $x$ and $y$ on the straight line joining them: it will be noticed that of the definitions given in this section, the last is the only one which is not valid in space of any number of dimensions. If $x$ is a point not on a straight line $h$, there are lines parallel to $h$ between $x$ and $h$, and therefore

3·64 $$(\exists h) . h \in \mathrm{Stl} . h \text{ part } (\iota`x,\ \Gamma) : \supset : (\exists h) . h \in \mathrm{Stl} . h \text{ part } (\iota`x,\ G`\Gamma),$$

and since the converse implication also is true we have

·65 $$(\exists h) . h \in \mathrm{Stl} . h \text{ part } (\iota`x,\ \Gamma) : \equiv : (\exists h) . h \in \mathrm{Stl} . h \text{ part } (\iota`x,\ G`\Gamma).$$

Following American writers, we describe a set as connected if in every genuine division into two parts one of the components contains a limiting point of the other:

·71 $$\mathrm{Cd} = \hat{\Gamma}\{\Gamma = \Theta \cup \Phi . \exists ! \Theta . \exists ! \Phi . \supset_{\Theta, \Phi} \exists ! \Theta \cap \Phi \cup \Theta \cap D`\Phi \cup D`\Theta \cap \Phi\} \quad \mathrm{Df.}$$

Substituting $\Gamma \cap \Delta$, $\Gamma \cap C`\Delta$ for $\Theta$, $\Phi$ in this definition we find

·72 $$\Gamma \in \mathrm{Cd} \supset : \exists ! \Gamma \cap \Delta . \exists ! \Gamma \cap C`\Delta . \supset_{\Delta} \exists ! \Gamma \cap B`\Delta :$$

a connected set cannot vault a boundary. On the other hand

·73 $$\Theta \cap G`\Phi = \Lambda . G`\Theta \cap \Phi = \Lambda . \Delta = \hat{x}\{(\exists y) . y \in \Theta . z \in \Phi \supset_z yz \geqslant 2xy\} . \supset : \Theta \subset \Delta . \Phi \subset C`\Delta . (\Theta \cup \Phi) \cap B`\Delta = \Lambda,$$

shewing that if a set is not connected there is a boundary which it does vault. Thus we have

·74 $$\mathrm{Cd} = \hat{\Gamma}\{\exists ! \Gamma \cap \Delta . \exists ! \Gamma \cap C`\Delta . \supset_{\Delta} \exists ! \Gamma \cap B`\Delta\},$$

the fundamental theorem that expresses the precise degree of continuity belonging to a connected set. From ·74 it follows that every united set is connected, a result of which the symbolical expression is

·75 $$\mathrm{Ud} \subset \mathrm{Cd} ;$$

the converse of ·75 is proved false by the actual construction of connected sets that are not united.

## 4. *Straight lines and rays.*

In the following discussion of certain parts of the theory of sets of points in a plane considerable use is made of sets of several particular kinds, which we commence by describing, and we reserve particular symbols for sets of these kinds.

To denote that a set $\Gamma$ consists of all the points composing a straight line we write $\Gamma \in \mathrm{Stl}$, and we use $g$, $h$, and $k$ only for straight lines:

4·11 $$\mathrm{Stl} = \text{straight line} \quad \mathrm{Df.}$$

If $h$ is any line through a point $x$, the set $h - \iota`x$ consists of two similar cells, one on each side of $x$; each of these cells is called in this paper* a ray, and of these rays $x$ is called the source and $h$ the line. For formal definition we take

·12 $$\mathrm{Ray} = \hat{\Gamma}\{(\exists x, h) . h \in \mathrm{Stl} . x \in h . \Gamma \in \kappa`(h - \iota`x)\} \quad \mathrm{Df},$$

* The most useful sense of the word ray in pure mathematics is to denote a directed straight line, but directed lines are not required in the study of fields and cordons, and the word ray is convenient; Hamilton, *Elements of Quaternions*, § 132, ex. 4 (p. 119 of the first (1866) edition; pp. 121, 122 of the first volume of Joly's (1899) edition), uses ray in precisely the sense adopted here.

and for rays we use $a$ and $c$; to denote that $x$ is the source of the ray $a$ we write $a \operatorname{ef} x$ or $x \operatorname{fe} a$, and to denote that $h$ is the line containing the ray $a$ we write $h \operatorname{lc} a$ or $a \operatorname{cl} h$, the use of any of these expressions being taken to imply that $a$ is a ray and also that $x$ or $h$ as the case may be is a point or line; a ray $a$ has only one source and only one line, and these may properly be denoted by $\operatorname{fe}‘a$ and $\operatorname{lc}‘a$, but the rays issuing from a common source $x$ form a class $(\operatorname{sg}‘\operatorname{ef})‘x$ and the rays situated in a line $h$ form a class $(\operatorname{sg}‘\operatorname{cl})‘h$ and both these classes have infinitely many members:

·13 $$\operatorname{lc} = \hat{h}\hat{a}\{a \in \operatorname{Ray} . h \in \operatorname{Stl} . a \subset h\} \quad \operatorname{Df},$$

·14 $$\operatorname{cl} = \operatorname{Cnv}‘\operatorname{lc} \quad \operatorname{Df},$$

·15 $$\operatorname{lc} \in 1 \rightarrow \operatorname{Cls},$$

·16 $$\operatorname{ef} = \hat{a}\hat{x}\{a \in \operatorname{Ray} . a \in \kappa‘(\operatorname{lc}‘a - \iota‘x)\} \quad \operatorname{Df},$$

·17 $$\operatorname{fe} = \operatorname{Cnv}‘\operatorname{ef} \quad \operatorname{Df},$$

·18 $$\operatorname{fe} \in 1 \rightarrow \operatorname{Cls}.$$

If $a$ is a ray, $\operatorname{lc}‘a - \iota‘\operatorname{fe}‘a$ consists of two cells, each of which is a ray; one of these rays is the ray $a$ itself, the other is called the reflex of $a$ and we denote it by $\operatorname{rfl}‘a$:

·21 $$\operatorname{rfl} = \hat{c}\hat{a}\{a \in \operatorname{Ray} . c = \operatorname{lc}‘a - \iota‘\operatorname{fe}‘a - a\} \quad \operatorname{Df},$$

·22 $$a \in \operatorname{Ray} \supset : \operatorname{E} ! \operatorname{rfl}‘a . \operatorname{rfl}‘a \in \operatorname{Ray}.$$

If $x$, $y$ are any two distinct points there is one and only one ray issuing from $x$ which contains $y$, and we denote this ray by $x \rightarrow y$; the reflex ray, which issues from the same source in the direction away from $y$, we denote by $x \leftarrow y$. The rays issuing from $x$ form the class $(\operatorname{sg}‘\operatorname{ef})‘x$, and the sets containing $y$ form the class $(\operatorname{gs}‘\epsilon)‘y$, but we must define $x \rightarrow y$ as $s‘\{(\operatorname{sg}‘\operatorname{ef})‘x \cap (\operatorname{gs}‘\epsilon)‘y\}$, or from a simpler formula by

·31 $$x \rightarrow y = s‘\hat{a}\{a \operatorname{ef} x . y \in a\} \quad \operatorname{Df},$$

not as $\imath‘\hat{a}\{a \operatorname{ef} x . y \in a\}$, for although

·32 $$x \neq y \supset \operatorname{E} ! \imath‘\hat{a}\{a \operatorname{ef} x . y \in a\},$$

and for any class of sets $\gamma$

·33 $$\operatorname{E} ! \imath‘\gamma \supset . \imath‘\gamma = s‘\gamma,$$

the class $\hat{a}\{a \operatorname{ef} x . x \in a\}$ is the null class of sets, and $\imath‘\hat{a}\{a \operatorname{ef} x . x \in a\}$ does not denote $\Lambda$ but is meaningless; on the other hand $s‘\hat{a}\{a \operatorname{ef} x . x \in a\}$ denotes by definition the set $\hat{z}\{(\exists a) a \operatorname{ef} x . x, z \in a\}$, and since the condition $(\exists a) . a \operatorname{ef} x . x, z \in a$ can in no way be satisfied, this set is the null set $\Lambda$. Thus ·31 yields as we desire

·34 $$x \neq y \supset : x \rightarrow y \in \operatorname{Ray} . x \rightarrow y \operatorname{ef} x . y \in x \rightarrow y,$$

·35 $$x \rightarrow x = \Lambda.$$

Considerations somewhat similar prevent us from defining $x \leftarrow y$ formally as $\operatorname{rfl}‘x \rightarrow y$: the null set is not a ray and this definition would leave $x \leftarrow x$ meaningless: it is sufficient to put

·36 $$x \leftarrow y = s‘\hat{a}\{a \operatorname{ef} x . y \in \operatorname{rfl}‘a\} \quad \operatorname{Df},$$

and then we have

4·37 $$x \neq y \supset : x \leftarrow y = \mathrm{rfl}`x \rightarrow y,$$

·38 $$x \leftarrow x = \Lambda.$$

The set $x \leftarrow y$ must not be confused with the set $y \rightarrow x$; both are null if $y$ coincides with $x$, but in general the former is a proper part of the latter:

·39 $$x \neq y \supset : x \leftarrow y \subset y \rightarrow x \,.\, \exists !\, y \rightarrow x - x \leftarrow y.$$

## 5. *Chords.*

If $y$ is distinct from $x$, the common part of the sets $x \rightarrow y$, $y \rightarrow x$ is the set formed of all points between $x$ and $y$ on the line through them; in any case this set is called the chord $xy$ and denoted by $x - y$:

5·11 $$x - y = x \rightarrow y \cap y \rightarrow x \quad \mathrm{Df};$$

the set obtained by adding to this chord the point $x$ is denoted by $x \vdash y$ or $y \dashv x$, and the set obtained by adding both the end points $x$, $y$ by $x \vdash\!\dashv y$:

·12 $$x \vdash y = x - y \cup \iota`x \quad \mathrm{Df},$$

·13 $$x \dashv y = x - y \cup \iota`y \quad \mathrm{Df},$$

·14 $$x \vdash\!\dashv y = x - y \cup \iota`x \cup \iota`y \quad \mathrm{Df}.$$

Two useful elementary propositions are

·21 $$z \in x \rightarrow y \supset . x \rightarrow z = x \rightarrow y,$$

·22 $$z \in x \leftarrow y \equiv x \in y - z,$$

of which the second is equivalent to

·23 $$x \leftarrow y = \hat{z}\{x \in y - z\};$$

and we use also

·24 $$y \neq x \,.\equiv x \in D`(x - y),$$

·25 $$y \neq x \,.\equiv x \in D`(x \vdash\!\dashv y).$$

If $y$ coincides with $x$, the chord $x - y$ is null, but the completed chord $x \vdash\!\dashv y$ has the one member $x$; the case of coincidence is the only case in which the derivative of the chord is contained in the chord, and also the only case in which the finished chord is not contained in its derivative:

·26 $$y = x \,.\equiv.\, x - y = \Lambda,$$

·27 $$y = x \,.\equiv.\, x \vdash\!\dashv y = \iota`x.$$

## 6. *Triangular domains.*

If three points $x$, $y$, $z$ are not collinear and $u$ is any point in the interior of the triangle of which they are the vertices, there is a length $\rho$, namely the length of the shortest, or of one of the shortest, of the perpendiculars from $u$ on the sides of the triangle, such that every point $v$ whose distance from $u$ is less than $\rho$ also lies in the interior of the triangle; in other words, the interior of the triangle is a domain. This domain is called the triangular domain *xyz*, and we denote it by tridom‘$(x, y, z)$; it can be defined formally in terms of chords, a simple though unsymmetrical definition being

6·11 $$\text{tridom‘}(x, y, z) = \hat{u}\,\{(\exists v) \,.\, v \in y - z \,.\, u \in x - v\} - (x - y) - (x - z) \quad \text{Df},$$

where the chords $x - y$, $x - z$, which in general have no points in common with the set $\hat{u}\,\{(\exists v) \,.\, v \in y - z \,.\, u \in x - v\}$, are subtracted in order that we may have

·12 $$(\exists h) \,.\, h \in \text{Stl} \,.\, x, y, z \in h \,.\supset \text{tridom‘}(x, y, z) = \Lambda,$$

an implication which can be replaced by the equivalence

·13 $$(\exists h) \,.\, h \in \text{Stl} \,.\, x, y, z \in h \,.\equiv.\, \text{tridom‘}(x, y, z) = \Lambda.$$

Since the framing of a definition more symmetrical in appearance than ·11 finds a natural place later in our work, we content ourselves for the present with ·11. Following a course which we take in a number of similar cases, we write

·14 $$\text{Tridom} = \hat{\Gamma}\,\{(\exists x, y, z) \,.\, \Gamma = \text{tridom‘}(x, y, z)\} \quad \text{Df},$$

and we must note the property implied in the name, expressed in the theorem

·15 $$\text{Tridom} \subset \text{Dom},$$

which is true even if the domain is in fact null.

## 7. *Parallel lines, and rays contained in parallel lines.*

To denote that two lines $h$, $k$ are parallel we write $h \,\text{prl}\, k$, it being understood that the possibility of coincidence is not excluded. Since our space is the Euclidean plane we can write

7·11 $$\text{prl} = \hat{h}\hat{k}\,\{h, k \in \text{Stl} : h = k \,.\vee.\, h \cap k = \Lambda\} \quad \text{Df},$$

but an interesting alternative rests on the fact that if $h$ and $k$ are not parallel they divide the plane into four pieces:

·12 $$h, k \in \text{Stl} \,.\supset:\ \text{Nc‘}\kappa\text{‘}\{C\text{‘}(h \cup k)\} = 2 \,.\equiv.\, h = k,$$

·13 $$h, k \in \text{Stl} \,.\supset:.\ \text{Nc‘}\kappa\text{‘}\{C\text{‘}(h \cup k)\} = 3 \,.\equiv:\, h \,\text{prl}\, k \,.\, h \neq k,$$

·14 $$h, k \in \text{Stl} \,.\supset:\ \text{Nc‘}\kappa\text{‘}\{C\text{‘}(h \cup k)\} = 4 \,.\equiv \sim h \,\text{prl}\, k.$$

Rays in parallel lines may have either opposite directions or a common direction. Utilising a simple criterion for two rays to have opposite directions we can take as definitions

·21 $$\text{opd} = \hat{a}\hat{c}\,\{a, c \in \text{Ray} \,.\, (\exists h, k, \Gamma, \Delta) \,.\, h, k \in \text{Stl} \,.\, \Gamma, \Delta \in \kappa\text{‘}C\text{‘}(a \cup c \cup \text{fe‘}a \mapsto \text{fe‘}c) \,.\, h \subset \Gamma \,.\, k \subset \Delta \,.\, \Gamma \neq \Delta\} \quad \text{Df},$$

·22 $$\text{cod} = \text{opd}^2 \quad \text{Df}.$$

## 8. *Leaves and clipped leaves.*

If $h$ is a line through a point $x$ and $a$, $c$ are the two rays forming $h - \iota\text{‘}x$, the sets $a$, $c$ are ordinally similar and so are the sets $G\text{‘}a$, $G\text{‘}c$, but since $a \cup c$ is not the whole of $h$ and one point of $h$ is contained both in $G\text{‘}a$ and in $G\text{‘}c$, neither $a$ nor $G\text{‘}a$ can properly be described as a half-line. Similarly if $h$ is any line, $C\text{‘}h$ is formed of two similar cells, but neither of these can be called a half-plane, and we call them clipped leaves, writing

8·11 $$\text{Clif} = \hat{\Gamma}\{(\exists h) . h \in \text{Stl} . \Gamma \in \kappa\text{‘}C\text{‘}h\} \quad \text{Df},$$

and using $\Upsilon$ and $\Omega$ for sets of this kind. If $\Upsilon$ is a clipped leaf, the complement of its boundary is composed of two cells each of which is a clipped leaf; one of these is $\Upsilon$ itself, the other we shall call the reflex of $\Upsilon$ and denote by rflx‘$\Upsilon$, but as we wish to postpone the formal definition we denote it for the present by $C$‘$G$‘$\Upsilon$, noting that

·12 $$\Upsilon \in \text{Clif} \supset C\text{‘}G\text{‘}\Upsilon \in \text{Clif}.$$

If $x$ is any point of the boundary of a clipped leaf $\Upsilon$, and if $a$ is one of the rays with source $x$ contained in $B$‘$\Upsilon$, the sets $\Upsilon \cup a$, $C$‘$G$‘$\Upsilon \cup$ rfl‘$a$ are similar mutually exclusive sets whose sum omits from the whole plane only the one point $x$, and the sets $\Upsilon \cup G$‘$a$, $C$‘$G$‘$\Upsilon \cup G$‘rfl‘$a$ obtained by adding to each of them the point $x$ are similar, their sum is the whole plane, and their only common point is $x$; neither $\Upsilon \cup a$ nor $\Upsilon \cup G$‘$a$ can be called a half-plane, but $\Upsilon \cup G$‘$a$ is a typical set of a kind of which we have to make much use, and we call such a set, that is, the set formed of a completed ray and all the points on one side of the line containing the ray, a leaf, and the source of the ray we call the pivot of the leaf. The definition of a leaf that follows explicitly the description just given is

·21 $$\text{Leaf} = \hat{\Gamma}\{(\exists a, \Delta) . \Delta \in \kappa\text{‘}C\text{‘lc‘}a . \Gamma = \iota\text{‘fe‘}a \cup a \cup \Delta\} \quad \text{Df},$$

but an adequate definition which formally is simpler could be derived from the theorem

·22 $$\text{Leaf} = \hat{\Gamma}\{(\exists a) . a \in \text{Ray} . Y\text{‘}C\text{‘}\Gamma = a . Y\text{‘}\Gamma = \text{lc‘}a - a\}.$$

For leaves we use M and N, and in virtue of ·22 we may take for the definition of the pivot of a leaf M

·23 $$\text{pvt} = \hat{x}\hat{\text{M}}\{\text{M} \in \text{leaf} . x = \text{fe‘}Y\text{‘}C\text{‘M}\} \quad \text{Df}.$$

Certain elementary properties of leaves have to be noted for use:

·31 $$\text{M} \in \text{Leaf} . x \in \text{M} . \supset : (\exists \text{N}) . \text{N} \in \text{Leaf} . x \,\text{pvt}\, \text{N} . \text{N} \subset \text{M},$$

·32 $$\text{M} \in \text{Leaf} . y, z \in C\text{‘M} . x \in y - z . \supset x \in C\text{‘M},$$

which implies

·33 $$\text{M} \in \text{Leaf} . x \in y - z \cap \text{M} . \supset : y \in \text{M} . \vee . z \in \text{M},$$

a result that proves valuable, and

·34 $$\text{M} \in \text{Leaf} . x \in \text{M} . \supset x \in D\text{‘}H\text{‘M},$$

which is used in conjunction with ·44 below.

Clipped leaves are to our main purpose of less importance than proper leaves, but they are simpler in nature, possessing the properties expressed by

·41 $$\text{Clif} \subset \text{Domex}$$

and by

·42 $$\Upsilon \in \text{Clif} . x \in \Upsilon . \supset : (\exists h) . h \in \text{Stl} . h \,\text{part}\, (\iota\text{‘}x, C\text{‘}\Upsilon).$$

But the results

·43 $$\Upsilon \in \text{Clif} . x \in \Upsilon . \supset : (\exists \text{M}) . \text{M} \in \text{Leaf} . x \in \text{M} . \text{M} \subset \Upsilon,$$

which can be strengthened into

$$\cdot 44 \qquad \Upsilon \epsilon \text{Clilf} . x \epsilon \Upsilon . \supset : (\exists M) . M \epsilon \text{Leaf} . x \epsilon H`M . M \subset \Upsilon,$$

and

$$\cdot 45 \qquad M \epsilon \text{Leaf} \supset H`M \epsilon \text{Clilf},$$

which is to some extent a converse of ·44, enable us often to secure the advantages of operating with clipped leaves.

## 9. *Sectors.*

If $a, c$ are two rays which have the same source $x$ but do not coincide, $C``(\iota`x \cup a \cup c)$ is the sum of two domains each of which has $\iota`x \cup a \cup c$ for its boundary and is called a sector of $a$ and $c$; if $\Delta$ is one of these sectors, $G`\Delta$ is obtained by adding the rays and the source to $\Delta$, and therefore the other sector is $C`G`\Delta$. When $c$ coincides with $a$, the set $C``(\iota`x \cup a \cup c)$ is a single domain, but it is convenient then to regard the null set as a sector of $a$ and $c$; in this case if $\Delta$ is the existent sector $C``(\iota`x \cup a)$, the completed set $G`\Delta$ is the whole plane, and $C`G`\Delta$ being null again represents the sector, although $C`G`\Lambda$ is not $\Delta$ but $G`\Delta$. For definition of sectors we can take

$$9\cdot 11 \qquad \text{Sectex} = \hat{\Gamma} \{(\exists x, a, c) . a, c \text{ ef } x . \Gamma \epsilon \kappa`C``(\iota`x \cup a \cup c)\} \quad \text{Df},$$

defining existent sectors, followed by

$$\cdot 12 \qquad \text{Sect} = \text{Sectex} \cup \iota`\Lambda \quad \text{Df}:$$

we reserve for sectors the letters $\Sigma$, T. The properties of sectors first to be noted are

$$\cdot 13 \qquad \text{Sectex} \subset \text{Domex},$$

implying

$$\cdot 14 \qquad \text{Sect} \subset \text{Dom},$$

and

$$\cdot 15 \qquad \Sigma \epsilon \text{Sectex} \supset C`G`\Sigma \epsilon \text{Sect}.$$

It is possible to replace ·12 or ·15 by

$$\cdot 16 \qquad \text{Sect} = \text{Sectex} \cup C``G``\text{Sectex},$$

and indeed to deal directly with sectors by starting from

$$\cdot 17 \qquad \text{Sect} = \hat{\Gamma} \{(\exists x, a, c, \Delta) : a, c \text{ ef } x . \Delta \epsilon \kappa`C``(\iota`x \cup a \cup c) : \Gamma = \Delta . \mathsf{v} . \Gamma = C`G`\Delta\},$$

which is effectively a combination of ·11 and ·16.

It is convenient to have symbolism expressing that $\Sigma$ is a sector, existent or null, of $a$ and $c$, but a direct construction is impeded by two considerations: unless $a$ and $c$ have a common source, the definition must not lead to the null set but is to fail altogether; nevertheless, the definition must depend primarily on the pair of rays, not on the sector, to meet the cases of the null sector and of the clipped leaf. To this end we write

$$\cdot 21 \qquad \text{sectex} = \hat{\Sigma}\hat{\gamma} \{(\exists x, a, c) : a, c \text{ ef } x . \gamma = \iota`a \cup \iota`c . \Sigma \epsilon \kappa`C``(\iota`x \cup a \cup c)\} \quad \text{Df},$$

$$\cdot 22 \quad \text{sect} = \hat{\Sigma}\hat{\gamma} \{(\exists x, a, c, \Delta) : a, c \text{ ef } x . \gamma = \iota`a \cup \iota`c . \Delta \epsilon \kappa`C``(\iota`x \cup a \cup c) : \Sigma = \Delta . \mathsf{v} . \Sigma = C`G`\Delta\} \quad \text{Df},$$

obtaining implicitly definitions of $\Sigma$ sectex$(\iota`a \cup \iota`c)$ and $\Sigma$ sect$(\iota`a \cup \iota`c)$; to avoid the use of the argument $\iota`a \cup \iota`c$ we substitute $(a, c)$, and we have

9·23 $$\Sigma \text{ sectex } (a, c) \equiv : (\exists x) \,.\, a, c \text{ ef } x \,.\, \Sigma \,\epsilon\, \kappa`C``(\iota`x \cup a \cup c),$$

·24 $$\Sigma \text{ sect } (a, c) \equiv : (\exists x, \Delta) : a, c \text{ ef } x \,.\, \Delta \,\epsilon\, \kappa`C``(\iota`x \cup a \cup c) : \Sigma = \Delta \,.\, \vee \,.\, \Sigma = C``G`\Delta.$$

Corresponding to ·12 is

·25 $$\Sigma \text{ sect } (a, c) \supset : \Sigma \text{ sectex } (a, c) \,.\, \vee \,.\, \Sigma = \Lambda,$$

but to obtain useful propositions we must exhibit the conditions under which the null set enters:

·26 $$\text{fe}`a = \text{fe}`c \,.\, a \neq c : \supset \,.\, \Sigma \text{ sect } (a, c) \equiv \Sigma \text{ sectex } (a, c),$$

·27 $$a = c \,.\, \supset : \text{E!}\, \text{sectex}`(a, c) \,.\, (\text{sg}`\text{sect})`(a, c) = \text{sectex}`(a, c) \cup \iota`\Lambda.$$

It is hardly necessary to add the propositions

·28 $$\text{Sectex} = \hat{\Sigma}\, \{(\exists a, c)\, \Sigma \text{ sectex } (a, c)\},$$

·29 $$\text{Sect} = \hat{\Sigma}\, \{(\exists a, c)\, \Sigma \text{ sect } (a, c)\}.$$

The relation to a sector $\Sigma$ of rays $a$, $c$ by means of which it is defined is expressed by calling these rays bounding rays of the sector, and we write

·31 $$\text{br} = \hat{a}\hat{\Sigma}\, \{(\exists c) \,.\, \Sigma \text{ sect } (a, c)\} \quad \text{Df},$$

·32 $$\text{rb} = \text{Cnv}`\text{br} \quad \text{Df}.$$

If $\Sigma$ is a sector of $a$ and $c$, the common source of $a$ and $c$ is called a vertex of $\Sigma$, and we put

·33 $$\text{vx} = \hat{x}\hat{\Sigma}\, \{(\exists a, c) \,.\, a, c \text{ ef } x \,.\, \Sigma \text{ sect } (a, c)\} \quad \text{Df},$$

·34 $$\text{xv} = \text{Cnv}`\text{vx} \quad \text{Df}.$$

If $x$ is a vertex of $\Sigma$ and a circle is described with centre $x$, the ratio of the length of the part of the circumference within $\Sigma$ to that of the whole circumference is called the angle of $\Sigma$ and denoted by ang$`\Sigma$; this angle is a definite one of the two angles between the bounding rays of $\Sigma$ which issue from $x$. If the angle of an existent sector is not equal to $\pi$, the vertex is unique and the bounding rays are definite, but if the angle is equal to $\pi$, the sector is a clipped leaf, every point of the boundary is a vertex, and every ray contained in the boundary is a bounding ray, peculiarities for which allowance has been made; to justify the use of the symbol ang$`\Sigma$ we have to remark that if the vertex is not unique the angle is the same at every vertex.

From the definition,

·41 $$\Sigma \text{ sectex } (a, c) \supset C``G`\Sigma \text{ sect } (a, c),$$

·42 $$\text{fe}`a = \text{fe}`c \,.\, \supset \,.\, \text{Nc}`(\text{sg}`\text{sect})`(a, c) = 2;$$

the sum of the angles of the two different sectors derived from one pair of conterminous rays is $2\pi$, and this is true if the two rays coincide, the null sector having angle 0 and a sector of the form $C``G`a$, where $a$ is a ray, having angle $2\pi$:

·43 $$(\exists a, c) \,.\, \Sigma, \text{T sect } (a, c) \,.\, \Sigma \neq \text{T} : \supset \text{ang}`\Sigma + \text{ang}`\text{T} = 2\pi.$$

The complement of a sector $\Sigma$ is not a sector, but is a completed sector or a completed ray according as $C`G`\Sigma$ exists or is null, and in particular the universal set $V$ is not a sector, the most comprehensive sector being the complement of a completed ray. It is hardly necessary to remark that in connection with some problems the valuable sets might be the complements of what we are calling sectors; in that case $V$ would be of the standard form while the null set would not. The propositions

·44 $$\Sigma \epsilon \text{Sectex} \supset \text{ang}`\Sigma \cup \text{ang}`C`G`\Sigma = 2\pi,$$

·45 $$\Sigma \epsilon \text{Sectex} \,.\, x \,\text{vx}\, \Sigma \,.\supset x \,\text{vx}\, C`G`\Sigma,$$

·46 $$\Sigma \epsilon \text{Sectex} \,.\, a \,\text{br}\, \Sigma \,.\supset a \,\text{br}\, C`G`\Sigma$$

are true even if the angle of $\Sigma$ is equal to $\pi$, for $\Sigma$ and $C`G`\Sigma$ acquire simultaneously the peculiarities consequent upon the possession of that special angle.

Of value to us in relation to any sector $\Sigma$ is the sector which we call the reflex of $\Sigma$, which may be described as the reflection of $\Sigma$ in a vertex of $\Sigma$; even if the angle of $\Sigma$ is equal to $\pi$, the reflex of $\Sigma$ is unique, for then the reflection of $\Sigma$ in each of its vertices is the same. We write

·51 $$\text{rflx} = \hat{\Gamma}\hat{\Sigma}\,[\Sigma \epsilon \text{Sect} \,.\, \Gamma = s`\hat{a}\,\{(\exists x) \,.\, x \,\text{vx}\, \Sigma \,.\, a \,\text{ef}\, x \,.\, \text{rfl}`a \subset \Sigma\}] \quad \text{Df},$$

to which an equivalent form is

·52 $$\text{rflx} = \hat{\Gamma}\hat{\Sigma}\,[\Sigma \epsilon \text{Sect} \,.\, \Gamma = \hat{z}\,\{(\exists x, y) \,.\, x \,\text{vx}\, \Sigma \,.\, y \epsilon \Sigma \,.\, x \epsilon y - z\}],$$

and we have

·53 $$\Sigma \epsilon \text{Sect} \supset \text{rflx}`\Sigma \epsilon \text{Sect},$$

·54 $$\text{ang}`\Sigma > \pi \equiv \exists \,!\, \Sigma \cap \text{rflx}`\Sigma,$$

·55 $$0 < \text{ang}`\Sigma < \pi \equiv \exists \,!\, C`G`\Sigma \cap \text{rflx}`C`G`\Sigma,$$

·56 $$\text{ang}`\Sigma \geqslant \pi \equiv C`G`\Sigma \subset \text{rflx}`\Sigma,$$

·57 $$\text{ang}`\Sigma \leqslant \pi \equiv \text{rflx}`\Sigma \subset C`G`\Sigma.$$

From the last two formulae,

·61 $$\text{ang}`\Sigma = \pi \,.\equiv.\, \text{rflx}`\Sigma = C`G`\Sigma,$$

and other distinctive properties of sectors of angle $\pi$ already mentioned are

·62 $$\text{ang}`\Sigma = \pi \,.\equiv \Sigma \epsilon \text{Clif},$$

and

·63 $$\Sigma \epsilon \text{Sectex} \supset : \text{ang}`\Sigma = \pi \,.\equiv.\, \text{Nc}`(\text{sg}`\text{vx})`\Sigma \neq 1,$$

·64 $$\text{ang}`\Sigma = \pi \,.\, \Sigma \,\text{Sect}\,(x \to y, x \to z) \,.\supset x \epsilon y - z,$$

which we require in the sequel; a slight but useful modification of ·56 is

·65 $$\text{ang}`\Sigma \geqslant \pi \,.\equiv.\, \Sigma \cup \text{reflx}`\Sigma = C``(\text{sg}`\text{vx})`\Sigma,$$

which involves

·66 $$\text{ang}`\Sigma > \pi \;\equiv.\, \text{Nc}`C``(\Sigma \cup \text{rflx}`\Sigma) = 1,$$

·67 $$\text{ang}`\Sigma = \pi \,.\equiv C``(\Sigma \cup \text{rflx}`\Sigma) \epsilon \text{Stl}.$$

## 10. *Circular domains.*

The last particular set which we have to mention is the circular domain, consisting in normal cases of all points inside a circle. Many of the properties of normal circular domains are shared by both the null set and the whole plane, and therefore we include these sets in the general definitions, which are

10·11 $$\text{Circdom} = \hat{\Gamma}\{(\exists x, \rho) . \Gamma = \hat{y}(xy < \rho)\} \quad \text{Df},$$

·12 $$\text{circdom}`(x, \rho) = \hat{y}\{xy < \rho\} \quad \text{Df},$$

the second, in which $\rho$ is assumed to be a signless number but not necessarily to be finite, defining the circular domain with centre $x$ and radius $\rho$; for circular domains we use $\Xi$, H, Z. With these definitions

·13 $$\text{Circdom} \subset \text{Dom},$$

·14 $$\Xi \in \text{Circdomex} \equiv : (\exists x, \rho) . \rho > 0 . \Xi = \hat{y}(xy < \rho),$$

existence being expressed in the usual manner; the only unlimited domain satisfying the condition* of ·11 is the plane itself, and therefore we can exclude this domain by considering the class Circdom $\cap$ Lm.

To indicate the relations between the circular domain circdom‘$(x, \rho)$ and the point $x$ and length $\rho$, we introduce the definitions

·21 $$\text{cent} = \hat{x}\hat{\Xi}\{(\exists \rho) . \Xi = \text{circdom}`(x, \rho)\} \quad \text{Df},$$

·22 $$\text{rad} = \hat{\rho}\hat{\Xi}\{(\exists x) . \Xi = \text{circdom}`(x, \rho)\} \quad \text{Df},$$

implying

·23 $$x \text{ cent } \Xi . \equiv : (\exists \rho) . \Xi = \text{circdom}`(x, \rho),$$

·24 $$\rho \text{ cent } \Xi . \equiv : (\exists x) . \Xi = \text{circdom}`(x, \rho).$$

A circular domain $\Xi$ has a unique radius, which can be denoted by rad‘$\Xi$; the centre is unique provided that the domain is neither the whole plane nor the null set:

·25 $$\text{rad} \in 1 \rightarrow \text{Cls},$$

·26† $$\text{cent} \restriction \text{Circdomex} \cap \text{Lm} \in 1 \rightarrow \text{Cls}.$$

The use to us of circular domains is in connection with limiting points of sets, for 3·22 is equivalent to each of the theorems

·31 $$x \in D`\Gamma . \equiv . \rho > 0 \supset_\rho \exists ! \Gamma \cap \{\text{circdom}`(x, \rho) - \iota`x\},$$

·32 $$x \in C`D`\Gamma . \equiv : (\exists \Xi) . x \text{ cent } \Xi . \Xi \subset C`\Gamma \cup \iota`x.$$

* A different order of ideas includes the clipped leaf as a form of unlimited circular domain, since the straight line is in one sense a form of circle; in that work however the distinction between the inside and the outside of a circle tends to lose importance, and the valuable construct is the cell of the complement of the circle, a set of which no use can be made in the study of limiting points.

† If $R$ is any relation, $R \restriction \beta$ denotes the same relation restricted in application to members of the class $\beta$, that is, denotes $\hat{x}\hat{y}(x R y . y \in \beta)$; similarly $\alpha \upharpoonleft R$, $\alpha \upharpoonleft R \restriction \beta$ denote $\hat{x}\hat{y}(x R y . x \in \alpha)$, $\hat{x}\hat{y}(x R y . x \in \alpha . y \in \beta)$. See *Principia Mathematica*, *35.

11. *Theorems relating to chords, leaves, sectors, and circular domains.*

Many of the properties of the special sets we have described are useful to us chiefly in the form of existence theorems. Thus we require

11·11 $x \epsilon y - z \,.\, y \,\mathrm{cent}\, \mathrm{H} \,.\, z \,\mathrm{cent}\, \mathrm{Z} \,.\, 0 < \mathrm{rad}\text{‘}\mathrm{H}, \mathrm{rad}\text{‘}\mathrm{Z} < \rho \,.$

$$\supset :. (\exists \Xi) : x \,\mathrm{cent}\, \Xi \,.\, 0 < \mathrm{rad}\text{‘}\Xi < \rho : v \epsilon \mathrm{H} \,.\, w \epsilon \mathrm{Z} \,.\, v \neq w \,.\, \supset_{v, w} \exists ! (v - w) \cap (\Xi - \iota\text{‘}x),$$

·12 $x \epsilon y - z \,.\, z \,\mathrm{cent}\, \mathrm{Z} \,.\, 0 < \mathrm{rad}\text{‘}\mathrm{Z} < \rho \,.$

$$\supset :. (\exists \Xi) : x \,\mathrm{cent}\, \Xi \,.\, 0 < \mathrm{rad}\text{‘}\Xi < \rho \,.\, w \epsilon \mathrm{Z} - \iota\text{‘}y \supset_w \exists ! (y - w) \cap \Xi,$$

·13 $x \epsilon y - z \,.\, y \,\mathrm{cent}\, \mathrm{H} \,.\, z \,\mathrm{cent}\, \mathrm{Z} \,.\, \exists ! \mathrm{H}, \mathrm{Z} \,.$

$$\supset :. (\exists \Xi) : x \,\mathrm{cent}\, \Xi \,.\, \exists ! \Xi : u \epsilon \Xi \supset_u (\exists v, w) \,.\, v \epsilon \mathrm{H} \,.\, w \epsilon \mathrm{Z} \,.\, u \epsilon v - w\,;$$

the result

·14 $x \,\mathrm{vx}\, \Sigma \,.\, x \,\mathrm{vx}\, \mathrm{T} \,.\, \Sigma \cup s\text{‘}\hat{a} \{a \,\mathrm{ef}\, x \,.\, a \,\mathrm{br}\, \Sigma\} \subset \mathrm{T} \,.\, \exists ! \Sigma \,.\, \supset \mathrm{Nc}\text{‘}\kappa\text{‘}(\mathrm{T} - \Sigma) = 2$

gives significance to

·15 $x \,\mathrm{vx}\, \Sigma \,.\, x \,\mathrm{vx}\, \mathrm{T} \,.\, \Sigma \cup s\text{‘}\hat{a} \{a \,\mathrm{ef}\, x \,.\, a \,\mathrm{br}\, \Sigma\} \subset \mathrm{T} \,.\, \exists ! \Sigma \,.\, y, z \epsilon \mathrm{T} - \Sigma \,.\, \mathrm{ang}\text{‘}\mathrm{T} \leqslant \pi \,.$

$$\supset : y \epsilon K\text{‘}_{\mathrm{T}-\Sigma} z \,.\, \vee \,.\, \exists ! y - z \cap \Sigma.$$

Of a different kind are

·21 $x \,\mathrm{vx}\, \Sigma \,.\, \tau < \mathrm{ang}\text{‘}\Sigma \,.\, \supset : (\exists \mathrm{T}) \,.\, \mathrm{ang}\text{‘}\mathrm{T} = \tau \,.\, x \,\mathrm{vx}\, \mathrm{T} \,.\, \mathrm{T} \subset \Sigma \,.\, \exists ! \Sigma - \mathrm{T},$

·22 $x \,\mathrm{vx}\, \Sigma \,.\, \mathrm{ang}\text{‘}\Sigma > \pi \,.\, \supset : (\exists \mathrm{M}) \,.\, \mathrm{M} \epsilon \mathrm{Leaf} \,.\, x \epsilon \mathrm{M} \,.\, \mathrm{M} \subset \Sigma \cup \iota\text{‘}x,$

·23 $\mathrm{ang}\text{‘}\Sigma = \pi \,.\, a \,\mathrm{br}\, \Sigma \,.\, \supset (\Sigma \cup a \cup \iota\text{‘}\mathrm{fe}\text{‘}a) \epsilon \mathrm{Leaf},$

·24 $\mathrm{ang}\text{‘}\Sigma \geqslant \pi \,.\, x \,\mathrm{vx}\, \Sigma \,.\, \supset : (\exists \Upsilon) \,.\, \Upsilon \epsilon \mathrm{Clilf} \,.\, x \epsilon B\text{‘}\Upsilon \,.\, \Upsilon \subset \Sigma,$

·25 $\mathrm{M} \epsilon \mathrm{Leaf} \,.\, x \epsilon \mathrm{M} \,.\, \supset : (\exists \Upsilon) \,.\, \Upsilon \epsilon \mathrm{Clilf} \,.\, x \epsilon B\text{‘}\Upsilon \,.\, \Upsilon \subset \mathrm{M},$

the last of which we use in the form

·26 $\mathrm{M} \epsilon \mathrm{Leaf} \,.\, x \epsilon \mathrm{M} \,.\, \supset : (\exists \Sigma) \,.\, \mathrm{ang}\text{‘}\Sigma = \pi \,.\, x \,\mathrm{vx}\, \Sigma \,.\, \Sigma \subset \mathrm{M}.$

By actual construction

·31 $\Upsilon, \Omega \epsilon \mathrm{Clilf} \,.\, B\text{‘}\Upsilon \cap B\text{‘}\Omega \epsilon 1 \,.\, x \epsilon \Upsilon \cap \Omega \,.\, a, c \,\mathrm{ef}\, x \,.\, a \,\mathrm{cod}\, B\text{‘}\Upsilon - G\text{‘}\Omega \,.\, c \,\mathrm{cod}\, B\text{‘}\Omega - G\text{‘}\Upsilon \,.$

$$\supset : (\exists \Sigma) \,.\, \Sigma \,\mathrm{sect}\, (a, c) \,.\, \mathrm{ang}\text{‘}\Sigma > \pi \,.\, \Sigma \subset \Upsilon \cup \Omega,$$

·32 $\Upsilon, \Omega \epsilon \mathrm{Clilf} \,.\, B\text{‘}\Upsilon \cap B\text{‘}\Omega \epsilon 1 \,.\, x \epsilon \Upsilon - \Omega \,.\, a \,\mathrm{ef}\, x \,.\, a \,\mathrm{cod}\, B\text{‘}\Upsilon - G\text{‘}\Omega \,.\, c = x \rightarrow \breve{\iota}\text{‘}(B\text{‘}\Upsilon \cap B\text{‘}\Omega) \,.$

$$\supset : (\exists \Sigma) \,.\, \Sigma \,\mathrm{sect}\, (a, c) \,\mathrm{ang}\text{‘}\Sigma > \pi \,.\, \Sigma \subset \Upsilon \cup \Omega\,;$$

·31 implies

·33 $\Upsilon, \Omega \epsilon \mathrm{Clilf} \,.\, B\text{‘}\Upsilon \cap B\text{‘}\Omega \epsilon 1 \,.\, x \epsilon \Upsilon \cap \Omega \,.\, \supset : (\exists \Sigma) \,.\, \mathrm{ang}\text{‘}\Sigma > \pi \,.\, x \,\mathrm{vx}\, \Sigma \,.\, \Sigma \subset \Upsilon \cup \Omega,$

·32 implies

·34 $\Upsilon, \Omega \epsilon \mathrm{Clilf} \,.\, B\text{‘}\Upsilon \cap B\text{‘}\Omega \epsilon 1 \,.\, x \epsilon \Upsilon - \Omega \,.\, \supset : (\exists \Sigma) \,.\, \mathrm{ang}\text{‘}\Sigma > \pi \,.\, x \,\mathrm{vx}\, \Sigma \,.\, \Sigma \subset \Upsilon \cup \Omega,$

and by an interchange of $\Upsilon$ and $\Omega$ implies also

·35 $\Upsilon, \Omega \epsilon \mathrm{Clilf} \,.\, B\text{‘}\Upsilon \cap B\text{‘}\Omega \epsilon 1 \,.\, x \epsilon \Omega - \Upsilon \,.\, \supset : (\exists \Sigma) \,.\, \mathrm{ang}\text{‘}\Sigma > \pi \,.\, x \,\mathrm{vx}\, \Sigma \,.\, \Sigma \subset \Upsilon \cup \Omega\,;$

and since

·36 $$\Upsilon \cup \Omega = (\Upsilon \cap \Omega) \cup (\Upsilon - \Omega) \cup (\Omega - \Upsilon)$$

we have from ·33, ·34, ·35

·37 $\Upsilon, \Omega \epsilon \mathrm{Clilf} \supset :: B\text{‘}\Upsilon \,\mathrm{prl}\, B\text{‘}\Omega \,.\, \vee :. x \epsilon \Upsilon \cup \Omega \supset : (\exists \Sigma) \,.\, \mathrm{ang}\text{‘}\Sigma > \pi \,.\, x \,\mathrm{vx}\, \Sigma \,.\, \Sigma \subset \Upsilon \cup \Omega.$

The last result implies

·38 $\Upsilon \epsilon \mathrm{Clilf} \,.\, \mathrm{ang}\text{‘}\Sigma > \pi \,.\, x \epsilon \Upsilon \,.\, \supset : (\exists \mathrm{T}) \,.\, \mathrm{ang}\text{‘}\mathrm{T} > \pi \,.\, x \,\mathrm{vx}\, \mathrm{T} \,.\, \mathrm{T} \subset \Upsilon \cup \Sigma,$

for among the clipped leaves contained in a sector whose angle is greater than $\pi$, some have boundaries not parallel to a particular line $B$‘$\Upsilon$.

Somewhat opposite in character is another result proved by construction. We have

11·41 $\text{ang}‘\Sigma > \pi \,.\, x \text{ vx } \Sigma \,.\, x \text{ cent } \Xi \,.\, \exists ! \Xi \,.\, h \in \text{Stl} \,.\, B‘\Sigma \cap B‘\Xi \subset h \,.\, x \in C‘h \,.\, \Upsilon = K‘(x, C‘h):$
$\supset : \Upsilon \in \text{Clilf} \,.\, x \in \Upsilon \,.\, \Upsilon \subset \Sigma \cup \Xi :$

if $\text{ang}‘\Sigma < 2\pi$ and $\Xi$ is limited, $h$ is determined by the condition of passing through both the points common to the circumference $B$‘$\Xi$ and the pair of completed rays $B$‘$\Sigma$, $x$ is necessarily outside $h$ because $\text{ang}‘\Sigma \neq \pi$, and the cell of $C$‘$h$ which contains $x$ is contained in $\Sigma \cup \Xi$ because $\text{ang}‘\Sigma > \pi$; if $\text{ang}‘\Sigma = 2\pi$, $B‘\Sigma \cap B‘\Xi$ is a single point and $h$ may be any line through this point except the line through $x$ itself; we make no use of the latitude allowed, for we require ·41 only for the sake of the existence theorem implied, namely

·42 $\text{ang}‘\Sigma > \pi \,.\, \text{E} ! \text{cent}‘\Xi \,.\, \text{cent}‘\Xi \text{ vx } \Sigma \,.\, \supset : (\exists \Upsilon) \,.\, \Upsilon \in \text{Clilf} \,.\, \text{cent}‘\Xi \in \Upsilon \,.\, \Upsilon \subset \Sigma \cup \Xi.$

## 12. *Primary chords of a set.*

A point is said to be on a primary chord, or simply on a chord, of a set $\Gamma$ if it lies between two points of $\Gamma$ on the line joining them, and we denote the set formed of points on the chords of $\Gamma$ by $S$‘$\Gamma$:

12·11 $S‘\Gamma = \hat{x}\{(\exists y, z) \,.\, y, z \in \Gamma \,.\, x \in y - z\}$ Df.

If three points of $\Gamma$ are collinear, the middle one is a member of both $\Gamma$ and $S$‘$\Gamma$, while if $\Gamma$ consists of only two points neither of these belongs to $S$‘$\Gamma$: there is no general relation of inclusion between $\Gamma$ and $S$‘$\Gamma$. Of more value for technical purposes than $S$‘$\Gamma$ is the set defined by

·12 $L‘\Gamma = \hat{x}\{(\exists y, z) \,.\, y, z \in \Gamma \,.\, x \in y \mapsto z\}$ Df,

which we call the set derived from $\Gamma$ by simple linkage, the principal advantage of this set being that from the definition

·13 $\Gamma \subset L‘\Gamma.$

The fundamental relation between the sets $S$‘$\Gamma$ and $L$‘$\Gamma$ is

·14 $L‘\Gamma = \Gamma \cup S‘\Gamma,$

which may be expressed in terms of operators only, in the form

·15 $L = I \cup S,$

$I$ being the operator of identity; but the elementary relation

·16 $S‘\Gamma \subset L‘\Gamma$

is often useful. Since $\Gamma$ and $S$‘$\Gamma$ are not mutually exclusive we cannot express $S$‘$\Gamma$ simply in terms of $\Gamma$ and $L$‘$\Gamma$, but we have

·17 $S‘\Gamma = \hat{x}\{x \in L‘(\Gamma - \iota‘x)\}.$

From ·13,

·21 $\exists ! \Gamma \supset \exists ! L‘\Gamma,$

so that indeed

·22 $\exists ! \Gamma \equiv \exists ! L‘\Gamma,$

which is equivalent to

·23 $\Gamma = \Lambda \,.\, \equiv \,.\, L‘\Gamma = \Lambda;$

this is a convenient point at which to note

·24 $$\Gamma \epsilon 1 \equiv L`\Gamma \epsilon 1,$$

·25 $$\Gamma \epsilon 1 \supset L`\Gamma = \Gamma.$$

As to the existence of $S`\Gamma$ we have

·26 $$\exists ! S`\Gamma \equiv : (\exists y, z) . y, z \epsilon \Gamma . y \neq z.$$

that is

·27 $$\mathrm{Nc}`\Gamma > 1 \equiv \exists ! S`\Gamma,$$

so that

·28 $$\Gamma \epsilon 0 \cup 1 \equiv . S`\Gamma = \Lambda.$$

From ·11 and 5·24

·31 $$\exists ! S`\Gamma \supset \Gamma \subset D`S`\Gamma ;$$

if $\Gamma$ is null it is contained in every set, and therefore

·32 $$\Gamma \epsilon 0 . \vee . \mathrm{Nc}`\Gamma > 1 : \supset \Gamma \subset D`S`\Gamma,$$

while

·33 $$\Gamma \epsilon 1 \supset \exists ! \Gamma - D`S`\Gamma$$

and so

·34 $$\Gamma \epsilon 1 \equiv \exists ! \Gamma - D`S`\Gamma,$$

·35 $$\mathrm{Nc}`\Gamma \neq 1 \equiv \Gamma \subset D`S`\Gamma.$$

From ·35 we have

·36 $$\mathrm{Nc}`\Gamma \neq 1 \supset D`\Gamma \subset D`S`\Gamma,$$

and since if $\Gamma$ is a unit set $D`\Gamma$ is contained in every set we can assert without hypothesis

·37 $$D`\Gamma \subset D`S`\Gamma,$$

so that from ·14

·38 $$D`L`\Gamma = D`S`\Gamma.$$

Of another kind are

·41 $$\mathrm{Nc}`\Gamma \neq 1 . \equiv L`\Gamma \subset D`L`\Gamma,$$

·42 $$S`\Gamma \subset D`S`\Gamma,$$

that is

·43 $$\mathrm{Nc}`\Gamma \neq 1 . \equiv L`\Gamma \epsilon \mathrm{Ds},$$

·44 $$S`\Gamma \epsilon \mathrm{Ds} :$$

both $S`\Gamma$ and $L`\Gamma$ are dense sets except when $\Gamma$ has but a single member, in which case $L`\Gamma$ also has one and only one member and is not dense, but $S`\Gamma$ is null and formally is dense.

By combining ·37 and ·38 with ·43 we have

·51 $$\mathrm{Nc}`\Gamma \neq 1 \supset \Gamma \cup D`\Gamma \cup S`\Gamma \subset D`S`\Gamma,$$

a result used later.

As we shall see from examples, neither $S`\Gamma$ nor $L`\Gamma$ need be complete, but from the elementary propositions

·61 $$(y - z) \cup (v - w) \cup (z - w) \epsilon \mathrm{Cg},$$

·62 $$(y \mapsto z) \cup (v \mapsto w) \cup (z \mapsto w) \epsilon \mathrm{Cl} \cap \mathrm{Cg}$$

we have

·63 $$S`\Gamma \epsilon \mathrm{Cg},$$

·64 $$L`\Gamma \epsilon \mathrm{Ud}.$$

From the preliminary propositions 11·11 and 11·13 we have immediately

12·65 $$S`D`\Gamma \subset D`S`\Gamma,$$

·66 $$S`(\Gamma - D`C`\Gamma) \subset S`\Gamma - D`C`S`\Gamma,$$

and ·14, ·37, ·65, and ·38 imply

·67 $$L`D`\Gamma \subset D`L`\Gamma$$

without hypothesis as to Nc‘Γ. We have sometimes to use

·68 $$\Gamma \subset \Delta \,.\supset.\, S`\Gamma \subset S`\Delta,$$

·69 $$\Gamma \subset \Delta \,.\supset.\, L`\Gamma \subset L`\Delta,$$

but we use them as a rule without explicit reference.

For a few purposes it is convenient to write

·71 $$S`(\Gamma, \Delta) = \hat{x}\{(\exists y, z) \,.\, y \in \Gamma \,.\, z \in \Delta \,.\, x \in y - z\} \quad \text{Df},$$

·72 $$L`(\Gamma, \Delta) = \hat{x}\{(\exists y, z) \,.\, y \in \Gamma \,.\, z \in \Delta \,.\, x \in y \mapsto z\} \quad \text{Df}.$$

with which notation

·73 $$S`(\Gamma, \Gamma) = S`\Gamma,$$

·74 $$L`(\Gamma, \Gamma) = L`\Gamma.$$

We have no need to enunciate results corresponding to all those given for $S`\Gamma$ and $L`\Gamma$, but we note that while the use of 11·11 gives information concerning the sets $S`(D`\Gamma, D`\Delta)$ and $L`(D`\Gamma, D`\Delta)$, by using 11·12 we can draw the conclusions

·75 $$S`(\Gamma, D`\Delta) \subset D`S`(\Gamma, \Delta),$$

·76 $$L`(\Gamma, D`\Delta) \subset D`L`(\Gamma, \Delta),$$

with the particular cases

·77 $$S`(\Gamma, D`\Gamma) \subset D`S`\Gamma,$$

·78 $$L`(\Gamma, D`\Gamma) \subset D`L`\Gamma.$$

An important relation between the set $L`(\Gamma, \Delta)$ and sets of the form $L`\Gamma$ is most simply written in the form

·81 $$L`(\Gamma \cup \Delta) = L`\Gamma \cup L`\Delta \cup L`(\Gamma, \Delta);$$

this is certainly redundant, for $\Gamma \cup \Delta$ is contained both in $L`\Gamma \cup L`\Delta$ and in $L`(\Gamma, \Delta)$, but it is the most useful form, and since even if we write

·82 $$L`(\Gamma \cup \Delta) = L`\Gamma \cup L`\Delta \cup S`(\Gamma, \Delta)$$

we are not secure against repetition, ·81 if replaced should yield only to

·83 $$L`(\Gamma \cup \Delta) = \Gamma \cup \Delta \cup S`\Gamma \cup S`\Delta \cup S`(\Gamma, \Delta).$$

The set $S`(\Gamma \cup \Delta)$ cannot be expressed by any formula similar to ·81, but

·84 $$S`(\Gamma \cup \Delta) - (\Gamma \cup \Delta) = S`\Gamma \cup S`\Delta \cup S`(\Gamma, \Delta) - (\Gamma \cup \Delta).$$

From ·81, ·41, ·67, and ·78 we have

·85 $$\Gamma \in 1 \,.\vee.\, L`G`\Gamma \subset D`L`\Gamma,$$

and ·16, ·85, ·38, and ·28 imply

·86 $$S`G`\Gamma \subset D`S`\Gamma,$$

although this cannot be deduced from ·84 without the help of ·35 and ·37.

### 13. *Secondary chords of a set.*

If $u$, $v$, $w$ are three points of a set $\Gamma$ no one of which lies on the primary chord joining the other two, the chord joining any one of the three to any point on the primary chord joining the other two is called a secondary chord of $\Gamma$, and we denote the set composed of points on secondary chords of $\Gamma$ by $T‘\Gamma$. This description is designed to bring $T‘\Gamma$ into relation with $S‘\Gamma$, but in fact a point is on a secondary chord of $\Gamma$ if it lies in a triangular domain whose vertices belong to $\Gamma$. By means of the operator $L$ we can give a symmetrical appearance to the definition of the triangular domain $xyz$, for

13·11 $$\text{tridom}‘(x, y, z) = L^2‘(\iota‘x \cup \iota‘y \cup \iota‘z) - L‘(\iota‘x \cup \iota‘y \cup \iota‘z),$$

and $T‘\Gamma$ is defined formally by

·12 $$T‘\Gamma = \hat{x}\{(\exists u, v, w) \,.\, u, v, w \in \Gamma \,.\, x \in \text{tridom}‘(u, v, w)\} \quad \text{Df.}$$

Since

·21 $$x, y, z \in \Gamma \supset \text{tridom}‘(x, y, z) \subset T‘\Gamma,$$

we have the important theorem

·22 $$T‘\Gamma \in \text{Dom},$$

implying

·23 $$\exists \,!\, T‘\Gamma \supset T‘\Gamma \in \text{Domex}.$$

It can easily be shewn that

·31 $$S^2‘\Gamma = S‘\Gamma \cup T‘\Gamma,$$

but $S‘\Gamma$ and $T‘\Gamma$ are not in general mutually exclusive, and indeed

·32 $$\text{Nc}‘\Gamma > 3 \,.\, \exists \,!\, T‘\Gamma \,.\supset \exists \,!\, S‘\Gamma \cap T‘\Gamma,$$

while on the other hand neither the set $S‘\Gamma - T‘\Gamma$ nor the set $T‘\Gamma - S‘\Gamma$ plays any part in the developments we make. Corresponding to 12·68 and 12·69 we have

·33 $$\Gamma \subset \Delta \supset T‘\Gamma \subset T‘\Delta.$$

The set $L^2‘\Gamma$, the set derived from $\Gamma$ by double linkage, is one of the most interesting of the sets connected with $\Gamma$, and the value of $T‘\Gamma$ is owing partly to the fact that a graphic analysis of $L^2‘\Gamma$, though not an analysis into mutually exclusive sets, is given by

·41 $$L^2‘\Gamma = \Gamma \cup S‘\Gamma \cup T‘\Gamma.$$

Proposition ·41 written in the form

·51 $$L^2‘\Gamma = L‘\Gamma \cup T‘\Gamma$$

has a curious result when taken in conjunction with the hypothesis that $\Gamma$ is connected or is the sum of two connected parts, which can be used in the form given by

·52 $$\Gamma = \Theta \cup \Phi \,.\, \Theta, \Phi \in \text{Cd} \,.$$
$$\equiv : u, v, w \in \Gamma \supset_{u,v,w} (\exists y, z, \Delta) \,.\, y, z \in \iota‘u \cup \iota‘v \cup \iota‘w \,.\, y \neq z \,.\, \Delta \in \text{Cd} \,.\, y, z \in \Delta \,.\, \Delta \subset \Gamma :$$

of any three points of $\Gamma$, two lie in a connected set contained in $\Gamma$. We have

·53 $$x \in \text{tridom}‘(t, y, z) \supset . \, K‘\{y, C‘(\iota‘x \cup x \leftarrow z \cup x \leftarrow t)\} \neq K‘\{z, C‘(\iota‘x \cup x \leftarrow z \cup x \leftarrow t)\}$$

which implies immediately

13·54 $x \in \text{tridom}`(t, y, z) . \Delta \in \text{Cd} . y, z \in \Delta \supset \exists ! \Delta \cap (\iota`x \cup x \leftarrow z \cup x \leftarrow t);$

also

·541 $\exists ! \Delta \cap \iota`x \equiv x \in \Delta,$

·542 $z \in \Delta . \exists ! \Delta \cap x \leftarrow z . \supset x \in S`\Delta,$

and in the notation of 12·7

·543 $\exists ! \Delta \cap x \leftarrow t \equiv x \in S`(\Delta, \iota`t)$

and therefore

·544 $\Delta \cup \iota`t \subset \Gamma . \exists ! \Delta \cap x \leftarrow t . \supset x \in S`\Gamma;$

from ·54, ·541, ·542, ·544

·55 $x \in \text{tridom}`(t, y, z) . \Delta \in \text{Cd} . y, z \in \Delta . \Delta \cup \iota`t \subset \Gamma . \supset x \in \Gamma \cup S`\Gamma,$

and ·55 with ·12 and ·52 gives

·56 $\Gamma = \Theta \cup \Phi . \Theta, \Phi \in \text{Cd} . x \in T`\Gamma . \supset x \in L`\Gamma,$

that is

·57 $\Gamma = \Theta \cup \Phi . \Theta, \Phi \in \text{Cd} . \supset T`\Gamma \subset L`\Gamma.$

From ·57 and ·51 comes

·58 $\Gamma = \Theta \cup \Phi . \Theta, \Phi \in \text{Cd} . \supset L^2`\Gamma = L`\Gamma,$

a result which we shall appreciate more fully when we are better acquainted with the set $L^2`\Gamma$; ·58 of course implies

·59 $\Gamma = \Theta \cup \Phi . \Theta, \Phi \in \text{Cd} . \supset : n \geqslant 1 \supset . L^n`\Gamma = L`\Gamma.$

## 14. *Cross points of a set.*

We call a point $x$ a cross point of $\Gamma$ if there are two chords of $\Gamma$ having $x$ for their only common point, and we denote the set of cross points of $\Gamma$ by $X`\Gamma$:

14·11 $X`\Gamma = \hat{x} \{(\exists t, u, v, w) . t, u, v, w \in \Gamma . \iota`x = (t - u) \cap (v - w)\}$ Df.

As with $S`\Gamma$ and $T`\Gamma$, so with $X`\Gamma$,

·21 $\Gamma \subset \Delta \supset X`\Gamma \subset X`\Delta;$

and we need hardly remark that

·22 $X`\Gamma \subset S`\Gamma.$

The fact that renders necessary the introduction of $X`\Gamma$ is that if $\Gamma$ is contained wholly in two intersecting lines the point of intersection does not belong to $T`\Gamma$ although it may belong to $X`\Gamma$; if however this point $x$ belongs to $X`\Gamma$ and $y$ is any point not on the lines containing $\Gamma$, then $x$ belongs not only to $X`(\Gamma \cup \iota`y)$ but also to $T`(\Gamma \cup \iota`y)$, and we have therefore

·23 $\exists ! X`\Gamma \supset : X`\Gamma \subset T`\Gamma . \vee . (\exists h, k) . h, k \in \text{Stl} . \Gamma \subset h \cup k.$

·24 $X`\Gamma - T`\Gamma \in 0 \cup 1.$

Of the sets connected with $X`\Gamma$ and $T`\Gamma$ it is actually $T`\Gamma \cup X`\Gamma$, which is of course identical with $T`\Gamma \cup (X`\Gamma - T`\Gamma)$, that plays the most prominent part in our work.

If $x$ is a cross point of Γ, and $t-u$, $v-w$ are chords of Γ which have $x$ for their only common point, then if $\rho$ is less than the length of each of the perpendiculars from $x$ on the chords $t-v$, $t-w$, $u-v$, $u-w$, every point $y$ distinct from $x$ whose distance from $x$ is less than $\rho$ belongs to one of the four triangular domains whose vertices are three of the points $t$, $u$, $v$, $w$; hence

·31 $$x \in X\text{‘}\Gamma \supset : (\exists \rho) . 0 < xy < \rho \supset_y y \in T\text{‘}\Gamma,$$

that is

·32 $$x \in X\text{‘}\Gamma \supset :. (\exists \rho) : \rho > 0 . xy < \rho \supset_y y \in T\text{‘}\Gamma \cup X\text{‘}\Gamma,$$

and so from 13·22

·33 $$\exists ! T\text{‘}\Gamma \supset T\text{‘}\Gamma \cup X\text{‘}\Gamma \in \text{Domex},$$

or since the null set is a domain

·34 $$T\text{‘}\Gamma \cup X\text{‘}\Gamma \in \text{Dom}.$$

## 15. *Convex sets.*

A set of points Γ is said to be convex when if two points $y$, $z$ belong to Γ every point of $y-z$ necessarily belongs to Γ; we write Γ ε Cvx to denote that Γ is convex, the formal definition being

15·11 $$\text{Cvx} = \hat{\Gamma}\{S\text{‘}\Gamma \subset \Gamma\} \quad \text{Df.}$$

Convex sets of points have many important properties, some of which we shall develop as we proceed. From 12·14 we have

·12 $$\Gamma \in \text{Cvx} \equiv . L\text{‘}\Gamma = \Gamma,$$

a relation often more useful than the fundamental one on which the definition is founded. Since ·11 with 12·68 implies

·13 $$\Gamma \in \text{Cvx} \supset S^2\text{‘}\Gamma \subset \Gamma,$$

we have from 13·31

·14 $$\Gamma \in \text{Cvx} \supset T\text{‘}\Gamma \subset \Gamma,$$

and so also, using 13·22,

·15 $$\Gamma \in \text{Cvx} \supset . T\text{‘}\Gamma \cup X\text{‘}\Gamma \subset \Gamma.$$

From 12·64 and ·12

·16 $$\text{Cvx} \subset \text{Ud},$$

a result which has its use in connection with the nature of the boundary of a convex set.

Possibilities in the relations to a set Γ of the sets $S$‘Γ, $L$‘Γ can be illustrated by means of lines, rays, and leaves. If Γ is a line or a ray, $S$‘Γ and $L$‘Γ both coincide with Γ, and Γ is convex, although if Γ is a ray it is not a complete set. If Γ is a leaf, $S$‘Γ consists of all the points of Γ except the pivot, and $L$‘Γ coincides with Γ; a leaf Γ is convex, although there is a whole ray Γ‘$C$‘Γ which consists of limiting points of Γ not belonging to Γ. We have already in 8·32 asserted implicitly that the complement of a leaf is a convex set.

### 16. *Excluding sectors and the excluding angle of a point.*

We say that a sector $\Sigma$ excludes a set $\Gamma$ if no points of $\Gamma$ lie within the sector, writing

16·11 $$\text{exsect} = \hat{\Sigma}\hat{\Gamma}\{\Sigma \epsilon \text{Sect} . \Sigma \cap \Gamma = \Lambda\} \quad \text{Df};$$

we have to remember that the boundary of a sector is not contained in the sector, so that $\Sigma$ exsect $\Gamma$ is not inconsistent with $(\exists x) . x \text{ vx } \Sigma . x \epsilon \Gamma$ or with $\exists ! \Gamma \cap B{}^{\backprime}\Sigma$, and also that the null sector has every vertex and excludes every set. From the last convention it follows that the class of numbers

$$\text{ang}{}^{\backprime\backprime}\hat{\Sigma}\{x \text{ vx } \Sigma . \Sigma \text{ exsect } \Gamma\}$$

contains the number 0, and from 11·21 it follows that this class is a stretch; because the bounding rays of a sector are contained in the complement of the sector, this stretch cannot have an upper limit which does not belong to it, that is to say, the stretch has a maximum, and we call this maximum the excluding angle of the point $x$ and the set $\Gamma$, writing

·12 $$\text{ea}{}^{\backprime}(x, \Gamma) = \text{max}{}^{\backprime}\text{ang}{}^{\backprime\backprime}\{(\text{sg}{}^{\backprime}\text{xv}){}^{\backprime}x \cap (\text{sg}{}^{\backprime}\text{exsect}){}^{\backprime}\Gamma\} \quad \text{Df}.$$

A sector of which $x$ is a vertex, which excludes $\Gamma$, and has $\text{ea}{}^{\backprime}(x, \Gamma)$ for its angle, we call a limiting excluding sector of $x$ and $\Gamma$, and we write

·13 $$\Sigma \text{ les } (x, \Gamma) = : \text{ang}{}^{\backprime}\Sigma = \text{ea}{}^{\backprime}(x, \Gamma) . x \text{ vx } \Sigma . \Gamma \cap \Sigma = \Lambda \quad \text{Df};$$

the class $(\text{sg}{}^{\backprime}\text{les}){}^{\backprime}(x, \Gamma)$ certainly exists, although if $\text{ea}{}^{\backprime}(x, \Gamma)$ is zero the members of the class are null sectors:

·14 $$\exists ! (\text{sg}{}^{\backprime}\text{les}){}^{\backprime}(x, \Gamma).$$

From the nature of a maximum and from ·13,

·21 $$\exists ! \Gamma - \iota{}^{\backprime}x . \Sigma \text{ les } (x, \Gamma) . a \text{ ef } x . a \text{ br } \Sigma . \supset : \text{T} \epsilon \text{Sect} . a \subset \text{T} . \supset_{\text{T}} \exists ! \Gamma \cap (\text{T} - \Sigma)$$

so that also

·22 $$\exists ! \Gamma - \iota{}^{\backprime}x . \Sigma \text{ les } (x, \Gamma) . a \text{ ef } x . a \text{ br } \Sigma .$$
$$\supset :. \exists ! \Gamma \cap a . \vee : 0 < \rho \leqslant 2\pi \supset_{\rho} (\exists \text{T}) . \text{ang}{}^{\backprime}\text{T} = \rho . a \text{ br T} . \exists ! \Gamma \cap (\text{T} - \Sigma);$$

·23 $$\exists ! \Gamma - \iota{}^{\backprime}x . \Sigma \text{ les } (x, \Gamma) . x \text{ vx T} . \Sigma \cup s{}^{\backprime}(\text{sg}{}^{\backprime}\text{br}){}^{\backprime}\Sigma \subset \text{T} . \exists ! \Sigma . \text{K} \epsilon \kappa{}^{\backprime}(\text{T} - \Sigma) . \supset \exists ! \Gamma \cap \text{K}$$

has value because its hypothesis includes that of 11·14, implying that $\text{T} - \Sigma$ has two cells;

·24 $$\exists ! \Gamma - \iota{}^{\backprime}x . \Sigma, \text{T les } (x, \Gamma) . \Sigma \neq \text{T} . \supset . \Sigma \cap \text{T} = \Lambda,$$

·25 $$\text{ang}{}^{\backprime}\Sigma = \text{ea}{}^{\backprime}(x, \Gamma) . \Sigma \text{ xv } x . \supset : \exists ! \Gamma \cap \Sigma . \vee . \Sigma \text{ les } (x, \Gamma).$$

The case in which $\Gamma$ has only the one point $x$ is peculiar; in this case every sector of which $x$ is a vertex excludes $\Gamma$, and $2\pi$, the greatest angle a sector can have, is the excluding angle:

·31 $$\text{ea}{}^{\backprime}(x, \iota{}^{\backprime}x) = 2\pi.$$

But $\text{V} - \iota{}^{\backprime}x$ is not a sector, and the limiting excluding sectors are the complements of completed rays issuing from $x$. If $y$ is any point other than $x$, the excluding angle of $y$ for $\iota`x$ is $2\pi$ but there is only one limiting excluding sector, the complement of $\iota`y \cup y \rightarrow x$:

·32 $$\text{ea}`(y, \iota`x) = 2\pi,$$

·33 $$y \neq x \,.\, \Sigma \text{ les } (y, \iota`x) \,.\, \supset \Sigma = C`\{\iota`y \cup y \rightarrow x\}.$$

It is convenient to note

·34 $$\text{ea}`(x, \Gamma) < 2\pi \supset \exists ! \, \Gamma - \iota`x.$$

The case in which the excluding angle of $x$ and $\Gamma$ is zero also is peculiar; the excluding sectors are the null sectors of which the various rays issuing from $x$ are the bounding rays, and every existent sector with vertex $x$ contains points of $\Gamma$:

·35 $$\text{ea}`(x, \Gamma) = 0 \,.\, \equiv : x \text{ vx } \Sigma \,.\, \exists ! \, \Sigma \,.\, \supset_\Sigma \exists ! \, \Gamma \cap \Sigma.$$

There may or may not be rays from $x$ which do not contain points of $\Gamma$, the existence of such rays being from our point of view irrelevant.

We have of course

·41 $$\Gamma \subset \Delta \,.\, \Sigma \text{ exsect } \Delta \,.\, \supset \Sigma \text{ exsect } \Gamma,$$

which implies

·42 $$\Gamma \subset \Delta \supset \text{ea}`(x, \Gamma) \geqslant \text{ea}`(x, \Delta).$$

A particular case of ·41 is

·43 $$\Sigma \text{ exsect } G`\Gamma \supset \Sigma \text{ exsect } \Gamma;$$

on the other hand, because $\Sigma$ is a domain, $D`C`\Sigma$ is contained in $C`\Sigma$, and therefore $G`C`\Sigma$ is identical with $C`\Sigma$; hence

·44 $$\Gamma \subset C`\Sigma \supset G`\Gamma \subset C`\Sigma$$

that is to say

·45 $$\Sigma \text{ exsect } \Gamma \supset \Sigma \text{ exsect } G`\Gamma,$$

which taken with ·43 gives

·46 $$\Sigma \text{ exsect } \Gamma \equiv \Sigma \text{ exsect } G`\Gamma,$$

and implies for all positions of $x$

·47 $$\text{ea}`(x, G`\Gamma) = \text{ea}`(x, \Gamma).$$

## 17. *The classification by means of excluding angles.*

Just as sectors fall into three classes, composed respectively of those whose angles are greater than $\pi$, those whose angles are equal to $\pi$, and those whose angles are less than $\pi$, and the properties of a member of one of these classes for the most part differ widely from those of a member of another, so each point of a plane falls with respect to any set $\Gamma$ into one of three classes according to the value of the excluding angle of $\Gamma$ for it. We write

17·11 $$U`\Gamma = \hat{x}\,\{\text{ea}`(x, \Gamma) > \pi\} \quad \text{Df},$$

·12 $$V`\Gamma = \hat{x}\,\{\text{ea}`(x, \Gamma) = \pi\} \quad \text{Df},$$

·13 $$W`\Gamma = \hat{x}\,\{\text{ea}`(x, \Gamma) < \pi\} \quad \text{Df};$$

of the three sets so defined it is the last which has the simplest and most important properties, but the three sets are studied together.

From 16·47 follow

17·21 $\Gamma \subset \Delta . \Delta \subset G'\Gamma . \supset U'\Delta = U'\Gamma,$

·22 $\Gamma \subset \Delta . \Delta \subset G'\Gamma . \supset V'\Delta = V'\Gamma,$

·23 $\Gamma \subset \Delta . \Delta \subset G'\Gamma . \supset W'\Delta = W'\Gamma,$

propositions which would justify us in studying the sets $U'\Gamma$, $V'\Gamma$, $W'\Gamma$ first on the hypothesis that $\Gamma$ is complete, a course which we do not actually take.

Since

·31 $x \epsilon y - z \supset \text{ea}'(x, \iota'y \cup \iota'z) = \pi,$

·32 $x \epsilon \text{tridom}'(u, v, w) \supset \text{ea}'(x, \iota'u \cup \iota'v \cup \iota'w) < \pi,$

we have from 16·42

·33 $x \epsilon S'\Gamma \supset \text{ea}'(x, \Gamma) \leqslant \pi,$

·34 $x \epsilon T'\Gamma \supset \text{ea}'(x, \Gamma) < \pi,$

that is,

·35 $S'\Gamma \subset V'\Gamma \cup W'\Gamma,$

·36 $T'\Gamma \subset W'\Gamma;$

also

·37 $x \epsilon X'\Gamma \supset \text{ea}'(x, \Gamma) < \pi,$

that is,

·38 $X'\Gamma \subset W'\Gamma,$

and therefore

·39 $T'\Gamma \cup X'\Gamma \subset W'\Gamma.$

From 16·24 we have

·41 $\exists ! \Gamma - \iota'x . \text{ea}'(x, \Gamma) > 0 . \supset \text{Nc}'(\text{sg}'\text{les})'(x, \Gamma) \leqslant 2\pi \div \text{ea}'(x, \Gamma)$

which has the corollaries

·42 $\exists ! \Gamma - \iota'x . \text{ea}'(x, \Gamma) > \pi . \supset \text{Nc}'(\text{sg}'\text{les})'(x, \Gamma) = 1,$

·43 $\text{ea}'(x, \Gamma) = \pi \supset : \text{Nc}'(\text{sg}'\text{les})'(x, \Gamma) = 1 . \vee . \text{Nc}'(\text{sg}'\text{les})'(x, \Gamma) = 2.$

A special case of the first of the corollaries just enunciated gives

·51 $\text{ea}'(x, \Gamma) = 2\pi \supset : (\exists a) . a \text{ ef } x . \Gamma \subset \iota'x \cup a,$

against which we put the converse

·52 $a \text{ ef } x . \Gamma \subset \iota'x \cup a . \supset \text{ea}'(x, \Gamma) = 2\pi,$

of which 16·31, 16·32 are particular cases; ·42 itself may be written in the form

·53 $x \epsilon U'\Gamma \supset : \text{Nc}'(\text{sg}'\text{les})'(x, \Gamma) = 1 . \vee . \Gamma = \iota'x.$

Corollary ·43 can be simplified, for

·54 $\text{ang}'\Sigma = \pi . \text{ang}'\mathrm{T} = \pi . x \text{ vx } \Sigma . x \text{ vx } \mathrm{T} . \supset : \exists ! \Sigma \cap \mathrm{T} . \vee . \mathrm{T} = \text{rflx}'\Sigma,$

and therefore

·55 $x \epsilon V'\Gamma . \text{Nc}'(\text{sg}'\text{les})'(x, \Gamma) = 2 . \supset : (\exists h) . h \epsilon \text{Stl} . \Gamma \subset h,$

and so from ·52, 9·64

·56 $x \epsilon V'\Gamma . \text{Nc}'(\text{sg}'\text{les})'(x, \Gamma) = 2 . \supset x \epsilon S'\Gamma;$

thus

·57 $x \epsilon V'\Gamma . \supset :. \text{Nc}'(\text{sg}'\text{les})'(x, \Gamma) = 1 .$

$\vee : x \epsilon S'\Gamma . (\exists h) . h \epsilon \text{Stl} . \Gamma \subset h . (\text{sg}'\text{les})'(x, \Gamma) = \kappa'C'h.$

It is worth while to notice that, if ea‘$(x, \Gamma)$ is equal to $\pi$, there is only one line which can be the boundary of a limiting excluding sector whether the number of such sectors is one or two:

·58 $$x \,\epsilon\, V\text{‘}\Gamma \supset . \text{Nc‘}B\text{‘‘(sg‘les)‘}(x, \Gamma) = 1.$$

One consequence of 16·25 is

·61 $$\Sigma \text{ les } (x, \Gamma) \supset : \exists ! \, \Gamma \cap \text{rflx‘}\Sigma . \vee . \text{ rflx‘}\Sigma \text{ les } (x, \Gamma);$$

writing this in the form

·62 $$\Sigma \text{ les } (x, \Gamma) \supset : \exists ! \, \Gamma \cap \text{rflx‘}\Sigma . \vee . \Gamma \subset C\text{‘}(\Sigma \cup \text{rflx‘}\Sigma)$$

we can apply 9·66 if ang‘$\Sigma$ is greater than $\pi$ and ·56 if ang‘$\Sigma$ is equal to $\pi$, and we have

·63 $$x \,\epsilon\, U\text{‘}\Gamma . \Sigma \text{ les } (x, \Gamma) . \supset : \exists ! \, \Gamma \cap \text{rflx‘}\Gamma . \vee . \Gamma = \iota\text{‘}x,$$

·64 $$x \,\epsilon\, V\text{‘}\Gamma . \Sigma \text{ les } (x, \Gamma) . \supset :. \exists ! \, \Gamma \cap \text{rflx‘}\Gamma . \vee : x \,\epsilon\, S\text{‘}\Gamma . \Gamma \subset B\text{‘}\Sigma;$$

the most interesting application of ·61 occurs when ea‘$(x, \Gamma)$ is less than $\pi$, but before proceeding to this application we complete the deductions which have to be made at the present stage from the hypothesis that ea‘$(x, \Gamma)$ is greater than or equal to $\pi$.

With regard to the set $U$‘$\Gamma$, we have only to point out that 11·42 implies

·71 $$x \,\epsilon\, U\text{‘}\Gamma \supset :. x \,\epsilon\, \Gamma \cup D\text{‘}\Gamma . \vee : (\exists \Upsilon) . \Upsilon \,\epsilon\, \text{Clilf} . x \,\epsilon\, \Upsilon . \Gamma \cap \Upsilon = \Lambda.$$

## 18. *Points for which the excluding angle is equal to $\pi$.*

If $\Sigma$ is a sector of angle $\pi$ and $x$ is a vertex of $\Sigma$, and if $w$ is any point not in $\Sigma$ or on its boundary, then in order that the ray issuing from $x$ in the direction of the ray through $w$ from a point $u$ distinct from $x$ and not contained in $\Sigma$ should neither contain $w$ nor lie in $\Sigma$ or $B$‘$\Sigma$, the point $u$ must lie either on the boundary $B$‘$\Sigma$ or in the strip bounded by $B$‘$\Sigma$ and the line through $w$ parallel to $B$‘$\Sigma$ and must not lie in $x \vdash w$; the region to which $u$ is thus restricted we denote temporarily by $P_x$‘$w$, the nature of $\Sigma$ and the conditions as to the positions of $x$ and $w$ being implied:

18·11 $$P_x\text{‘}w = \hat{u}\,\{\text{ang‘}\Sigma = \pi . x \text{ vx } \Sigma . w \,\epsilon\, C\text{‘}G\text{‘}\Sigma . u \,\epsilon\, C\text{‘}\Sigma : x \rightarrow y \text{ cod } u \rightarrow w \supset_y y \,\epsilon\, C\text{‘}G\text{‘}\Sigma - \iota\text{‘}w\} \quad \text{Dft [18]},$$

·12 $$P_x\text{‘}w = s\text{‘}\hat{h}\,\{h \text{ prl } B\text{‘}\Sigma . \exists ! \, h \cap x \vdash w\} - x \vdash w.$$

If $u$ belongs to $P_x$‘$w$, there are four sectors which have for one bounding ray the ray from $x$ in the direction of $u \rightarrow w$ and for the other bounding ray a bounding ray of $\Sigma$, and of these four there is one and only one which contains $\Sigma$ and does not include $u$; this sector we denote for a time by $R_x$‘$u$:

·13 $$R_x\text{‘}u = \iota\text{‘}\hat{\mathrm{T}}\,\{u \,\epsilon\, P_x\text{‘}w . (\exists a, c) . a, c \text{ ef } x . a \text{ cod } u \rightarrow w . c \text{ br } \Sigma . a, c \text{ br } \mathrm{T} . \Sigma \subset \mathrm{T} . u \,\epsilon\, C\text{‘}\mathrm{T}\} \quad \text{Dft [18]}.$$

The properties of sectors of the form $R_x$‘$u$ relevant to our purpose are only two, namely

·14 $$\exists ! \, R_x\text{‘}u \supset . (\exists a) . a \text{ ef } x . a \text{ br } \Sigma . a \subset R_x\text{‘}u,$$

·15 $$v \,\epsilon\, R_x\text{‘}u - \Sigma \supset \exists ! \, u - v \cap x \vdash w.$$

The first of these, with 16·21, gives

18·21 $$x \epsilon V`\Gamma \,.\, \Sigma \text{ les } (x, \Gamma) \,.\, u \epsilon P_x`w \,.\, \supset \exists \,!\, \Gamma \cap R_x`u - \Sigma,$$

and so from the second we have

·22 $$x \epsilon V`\Gamma \,.\, \Sigma \text{ les } (x, \Gamma) \,.\, \exists \,!\, \Gamma \cap P_x`w \,.\, \supset \exists \,!\, S`\Gamma \cap x \vdash w.$$

We have no reason to suppose that for every position of $w$ in $C`G`\Sigma$ points of $\Gamma$ are to be found in $P_x`w$, and the alternatives we consider are

hyp 18 $a$ $(\exists z) : z \epsilon C`G`\Sigma \,.\, w \epsilon x - z \supset_w \exists \,!\, \Gamma \cap P_x`w,$

hyp 18 $b$ $(\exists y, z, v, w) : y, z \epsilon C`G`\Sigma \,.\, x \to y \neq x \to z \,.\, v \epsilon x - y \,.\, w \epsilon x - z \,.$
$G`P_x`v = G`P_x`w \,.\, \Gamma \cap P_x`v = \Lambda \,.\, \Gamma \cap P_x`w = \Lambda;$

the form adopted for the second of these is designed to shew that one of the hypotheses is necessarily fulfilled, but this second assumption is equivalent simply to

hyp 18 $b$ $(\exists v, w) : v, w \epsilon C`G`\Sigma \,.\, x \to v \neq x \to w \,.\, G`P_x`v = G`P_x`w \,.\, \Gamma \cap (P_x`v \cup P_x`w) = \Lambda;$

in both forms, the condition $G`P_x`v = G`P_x`w$ is a method that happens to be simple notationally of expressing that the line through $v$ and $w$ is parallel to $B`\Sigma$, a condition required in ·32 to ensure that $x - v$ is contained in $P_x`w$ and $x - w$ in $P_x`v$. From ·22 and 12·42 we have

·31 $$x \epsilon V`\Gamma \,.\, \Sigma \text{ les } (x, \Gamma) \,.\, \text{hyp } 18\,a \,.\, \supset x \epsilon D`S`\Gamma;$$

on the other hand

·32 $$x \epsilon V`\Gamma \,.\, \Sigma \text{ les } (x, \Gamma) \,.\, v, w \epsilon C`G`\Sigma \,.\, x \to v \neq x \to w \,.\, G`P_x`v = G`P_x`w \,.\, \supset \,.\, \Sigma \cup \iota`x \cup P_x`v \cup P_x`w \epsilon \text{ Clilf},$$

·33 $$\Gamma \cap \Sigma = \Lambda \,.\, \Gamma \cap (P_x`v \cup P_x`w) = \Lambda \,.\, \supset : x \epsilon \Gamma \,.\, \vee \,.\, \Gamma \cap (\Sigma \cup \iota`x \cup P_x`v \cup P_x`w) = \Lambda,$$

and therefore

·34 $$x \epsilon V`\Gamma \,.\, \Sigma \text{ les } (x, \Gamma) \,.\, \text{hyp } 18\,b \,.\, \supset : x \epsilon \Gamma \,.\, \vee \,.\, (\exists \Upsilon) \,.\, \Upsilon \epsilon \text{ Clilf} \,.\, x \epsilon \Upsilon \,.\, \Gamma \cap \Upsilon = \Lambda;$$

·31 and ·34 imply

·35 $$x \epsilon V`\Gamma \supset : x \epsilon D`S`\Gamma \,.\, \vee \,.\, (\exists \Upsilon) \,.\, \Upsilon \epsilon \text{ Clilf} \,.\, x \epsilon \Upsilon \,.\, \Gamma \cap \Upsilon = \Lambda,$$

since $\Gamma$ is contained in $D`S`\Gamma$ if $\Gamma$ has more than one point and $V`\Gamma$ is null in the case excepted.

## 19. *Points for which the excluding angle is less than $\pi$.*

Turning to the set $W`\Gamma$, we have first to conduct an investigation in some respects analogous to that leading to 18·35, but with a result ultimately of more value.

An immediate deduction from 16·22 is

19·11 $$0 < \text{ea}`(x, \Gamma) < \pi \,.\, \Sigma \text{ les } (x, \Gamma) \,.\, \text{rflx}`\Sigma \text{ les } (x, \Gamma) \,.\, \supset : x \epsilon X`\Gamma \,.\, \vee \,.\, \exists \,!\, \Gamma \cap C`G`(\Sigma \cup \text{rflx}`\Sigma),$$

and since we do not need to examine in the present connection the case of a cross point, the two sets of hypotheses which we consider in detail are

hyp 19 $a$ $\text{ang}`\Sigma < \pi \,.\, \Sigma \text{ les } (x, \Gamma) \,.\, \exists \,!\, \Gamma \cap \text{rflx}`\Sigma,$

hyp 19 $b$ $\text{ang}`\Sigma < \pi \,.\, \Sigma \text{ les } (x, \Gamma) \,.\, \text{rflx}`\Sigma \text{ les } (x, \Gamma) \,.\, \Sigma \text{ sect } (a, c) \,.$
$\exists \,!\, \Gamma \cap C`G`(\Sigma \cup \text{rflx}`\Sigma \cup a \cup c),$

which we distinguish as hyp 19 $a$ and hyp 19 $b$, writing $(\exists\Sigma)$ hyp 19 $a$ if there is a limiting excluding sector such that the first hypothesis is satisfied, $(\exists\Sigma)$ hyp 19 $b$ if there is one satisfying the second. From 17·62 and ·11

·12 $\quad \text{ea}`(x, \Gamma) < \pi \equiv : (\exists\Sigma) \text{ hyp } 19\, a \,.\, \mathbf{v} \,.\, (\exists\Sigma) \text{ hyp } 19\, b \,.\, \mathbf{v} \,.\, x \,\epsilon\, X`\Gamma,$

and though the possibilities are not mutually exclusive we can extract the information we require by treating them separately. We notice that

·13 $\quad \exists \,!\, \text{rflx}`\Sigma \supset 0 < \text{ang}`\Sigma,$

so that in hyp 19 $a$ we have actually $0 < \text{ang}`\Sigma < \pi$, and that in both cases, from 16·34, $\exists \,!\, \Gamma - \iota`x$.

The form of the last constituent of hyp 19 $b$ is designed to admit the possibility of null sectors; if $\Sigma$ is not null, the whole effect of adding to $\Sigma$ and rflx'$\Sigma$ their bounding rays and their vertices is to complete $\Sigma \cup \text{rflx}`\Sigma$, but if $\Sigma$ is a null sector no bounding ray consists of limiting points of $\Sigma$, and indeed $G`(\Sigma \cup \text{rflx}`\Sigma)$ as well as $\Sigma$ itself is null.

If $\Sigma$ is a sector with angle between 0 and $\pi$ and vertex $x$, then in order that the reflex of a ray $x \rightarrow y$ may be neither a part of $\Sigma$ or rflx'$\Sigma$ nor a bounding ray of $\Sigma$, the point $y$ must lie outside both $\Sigma$ and rflx'$\Sigma$ and must not belong to the boundary of rflx'$\Sigma$; $y$ must belong to $C`(\Sigma \cup G`\text{rflx}`\Sigma)$. The constructions we have to make in relation to a point $y$ which when ang'$\Sigma$ is between 0 and $\pi$ require $y$ to belong to $C`(\Sigma \cup G`\text{rflx}`\Sigma)$, we can make if $\Sigma$ is a null sector provided then that $y$ does not lie in the line containing the bounding ray of $\Sigma$. The regions concerned in the two cases are covered by the one definition

·21 $\quad P_x`\Sigma = \hat{y}\,\{a, c \text{ ef } x \,.\, \Sigma \text{ sect}\,(a, c) \,.\, x \leftarrow y \subset C`(\Sigma \cup \text{rflx}`\Sigma \cup a \cup c)\}$ Dft [19];

if ang'$\Sigma$ is not less than $\pi$, then $C`(\Sigma \cup \text{rflx}`\Sigma \cup a \cup c)$ consists of the one point $x$ and cannot contain any rays: hence $y \,\epsilon\, P_x`\Sigma$ is false unless ang'$\Sigma$ is less than $\pi$, and it is superfluous to introduce the condition $\text{ang}`\Sigma < \pi$ explicitly into the definition. If $y$ belongs to $P_x`\Sigma$, then the ray $x \leftarrow y$ is not a bounding ray of $\Sigma$, and of the sectors which have $x \leftarrow y$ for one bounding ray and a bounding ray of $\Sigma$ for the other bounding ray, there is one and only one which contains $\Sigma$ and does not contain $y$; this sector we denote temporarily by $R_x`(y, \Sigma)$, implying by the use of $R_x`(y, \Sigma)$ that $y$ lies in $P_x`\Sigma$:

·22 $\quad R_x`(y, \Sigma) = \check{\iota}`\hat{\text{T}}\,\{(\exists a, c) \,.\, a, c \text{ ef } x \,.\, \Sigma \text{ sect}\,(a, c) \,.\, y \,\epsilon\, P_x`\Sigma \,.\, \text{T sect}\,(a, x \leftarrow y) \,.$
$\Sigma \subset \text{T} \,.\, y \,\epsilon\, C`\text{T}\}$ Dft [19].

The use here of sectors of the form $R_x`(y, \Sigma)$ depends on the propositions

·23 $\quad \text{ea}`(x, \Gamma) < \pi \,.\, \Sigma \text{ les}\,(x, \Gamma) \,.\, \supset \exists \,!\, \Gamma \cap P_x`\Sigma,$

·24 $\quad \text{ea}`(x, \Gamma) < \pi \,.\, \Sigma \text{ les}\,(x, \Gamma) \,.\, \supset : (y) \,.\, \exists \,!\, \Gamma \cap \{R_x`(y, \Sigma) - \Sigma\},$

consequences of 16·21 and 16·35; these imply

·25 $\quad \text{hyp } 19\, a \;\supset : (\exists u, v, w) \,.\, u, v, w \,\epsilon\, \Gamma \,.\, u \,\epsilon\, \text{rflx}`\Sigma \,.\, w \,\epsilon\, R_x`(v, \Sigma) - \Sigma,$

·26 $\quad \text{hyp } 19\, b \;\supset : (\exists u, v, w) \,.\, u, v, w \,\epsilon\, \Gamma \,.\, v \,\epsilon\, R_x`(u, \Sigma) - \Sigma \,.\, w \,\epsilon\, R_x`(u, \text{rflx}`\Sigma) - \text{rflx}`\Sigma,$

and shew the kinds of properties of the sectors of the form $R_x`(y, \Sigma)$ that concern us.

From the definition

19·31 $$R_x\text{‘}(y, \Sigma) - \Sigma \subset P_x\text{‘}\Sigma$$

for all positions of $y$ in $P_x\text{‘}\Sigma$, and therefore

·32 $$z \in R_x\text{‘}(y, \Sigma) - \Sigma \supset \exists ! R_x\text{‘}(z, \Sigma);$$

moreover

·33 $$z \in R_x\text{‘}(y, \Sigma) - \Sigma \supset y \in R_x\text{‘}(z, \Sigma) - \Sigma,$$

so that

·34 $$z \in R_x\text{‘}(y, \Sigma) - \Sigma \equiv y \in R_x\text{‘}(z, \Sigma) - \Sigma$$

and the relation $z \in R_x\text{‘}(y, \Sigma) - \Sigma$, which ·25 and ·26 shew to be connected with the use of excluding angles, is symmetrical between $y$ and $z$.

We are now in immediate touch with the proposition we wish to establish, for

·41 $$u \in \text{rflx}\text{‘}\Sigma \,.\, w \in R_x\text{‘}(v, \Sigma) - \Sigma \,.\supset x \in \text{tridom}\text{‘}(u, v, w),$$

in which ang‘$\Sigma$ cannot be 0, and

·42 $$v \in R_x\text{‘}(u, \Sigma) - \Sigma \,.\, w \in R_x\text{‘}(u, \text{rflx}\text{‘}\Sigma) - \text{rflx}\text{‘}\Sigma \,.\supset x \in \text{tridom}\text{‘}(u, v, w);$$

and from these with ·12, ·25 and ·26, we have

·43 $$\text{ea}\text{‘}(x, \Gamma) < \pi \supset x \in T\text{‘}\Gamma \cup X\text{‘}\Gamma,$$

that is

·44 $$W\text{‘}\Gamma \subset T\text{‘}\Gamma \cup X\text{‘}\Gamma,$$

which taken with 17·39 gives

·45 $$T\text{‘}\Gamma \cup X\text{‘}\Gamma = W\text{‘}\Gamma,$$

and implies, from 14·33,

·46 $$W\text{‘}\Gamma \in \text{Dom}.$$

The last property of $W$‘$\Gamma$ which we wish to mention is deducible immediately from 16·23, 11·14 and 11·15:

·51 $$x \in W\text{‘}\Gamma \,.\, \Sigma \text{ les}(x, \Gamma) \,.\, \exists ! \Sigma \,.\supset \exists ! S\text{‘}\Gamma \cap \Sigma;$$

this result emphasises the possible discontinuity of the excluding angle regarded as a function of the position of $x$, for it gives immediately

·52 $$\Gamma \in \text{Cvx} \,.\, x \in W\text{‘}\Gamma \,.\, \Sigma \text{ les}(x, \Gamma) \,.\supset \Sigma = \Lambda,$$

that is

·53 $$\Gamma \in \text{Cvx} \supset : \text{ea}\text{‘}(x, \Gamma) < \pi \,.\equiv .\, \text{ea}\text{‘}(x, \Gamma) = 0.$$

## 20. *Excluding leaves and the points outside the field of a set.*

We say that a leaf M excludes a set $\Gamma$ if no points of $\Gamma$ belong to the leaf, and we write

20·11 $$\text{exlf} = \hat{\text{M}}\hat{\Gamma}\{\text{M} \in \text{Leaf} \,.\, \Gamma \cap \text{M} = \Lambda\} \quad \text{Df},$$

from which we have at once

·12 $$\Gamma \subset \Delta \,.\, \text{M exlf}\,\Delta \,.\supset \text{M exlf}\,\Gamma,$$

·13 $$\text{M} \in \text{Leaf} \,.\, \text{M} \subset \text{N} \,.\, \text{N exlf}\,\Gamma \,.\supset \text{M exlf}\,\Gamma;$$

the leaves excluding a set form a class (sg‘exlf)‘Γ which may be null but otherwise contains an infinity of numbers; from ·12

·14 $\Gamma \subset \Delta \supset (\mathrm{sg}‘\mathrm{exlf})‘\Delta \subset (\mathrm{sg}‘\mathrm{exlf})‘\Gamma$.

We say that a point is outside the field of a set if there is a leaf which contains the point and excludes the set; the points outside the field of a set Γ compose a set which we denote by $E‘\Gamma$:

·21 $E‘\Gamma = \hat{x}\{(\exists \mathrm{M}) . \mathrm{M}\ \mathrm{exlf}\ \Gamma . x \in \mathrm{M}\}$ Df,

which is equivalent to

·22 $E‘\Gamma = s‘(\mathrm{sg}‘\mathrm{exlf})‘\Gamma$.

From ·21, 8·31 and ·13 comes

·23 $x \in E‘\Gamma \supset : (\exists \mathrm{N}) . \mathrm{N}\ \mathrm{exlf}\ \Gamma . x\ \mathrm{pvt}\ \mathrm{N}$;

on the other hand

·24 $x\ \mathrm{pvt}\ \mathrm{N} \supset x \in \mathrm{N}$,

and therefore

·25 $\mathrm{N}\ \mathrm{exlf}\ \Gamma . x\ \mathrm{pvt}\ \mathrm{N} . \supset x \in E‘\Gamma$,

so that

·26 $E‘\Gamma = \hat{x}\{(\exists \mathrm{M}) . \mathrm{M}\ \mathrm{exlf}\ \Gamma . x\ \mathrm{pvt}\ \Gamma\}$,

or more compactly

·27 $E‘\Gamma = \mathrm{pvt}‘‘(\mathrm{sg}‘\mathrm{exlf})‘\Gamma$,

an elegant but not as far as we have found a useful result.

## 21. *The field and the cordon of a set.*

It is the set complementary to $E‘\Gamma$ which is of value in analysis: indeed, it is the known importance of this set, which we call the field of Γ, that justifies our whole study; we write

21·11 $F‘\Gamma = C‘E‘\Gamma$ Df,

but sometimes we make use of the equivalent

·12 $E‘\Gamma = C‘F‘\Gamma$.

A direct definition of $F‘\Gamma$ is of course

·13 $F‘\Gamma = \hat{x}\{\mathrm{M} \in \mathrm{Leaf} . x \in \mathrm{M} . \supset_{\mathrm{M}} \exists ! \Gamma \cap \mathrm{M}\}$,

but it is usually easier to deal with $E‘\Gamma$ defined by 20·21 than with $F‘\Gamma$ defined by ·13. The boundary of the field of Γ we call the cordon of Γ and denote by $Q‘\Gamma$; thus

·14 $Q‘\Gamma = B‘F‘\Gamma$ Df,

·15 $Q‘\Gamma = B‘E‘\Gamma$,

·16 $Q‘\Gamma = F‘\Gamma \cap D‘E‘\Gamma \cup E‘\Gamma \cap D‘F‘\Gamma$,

the last embodying the definition of the boundary. Every set possesses both a field and a cordon, but it is not every set that can serve in either of these capacities, for the fact of being a field or a cordon itself implies properties; it is convenient to write

·17 $\mathrm{Fld} = \hat{\Gamma}\{(\exists \Delta) . \Gamma = F‘\Delta\}$ Dft [21–26],

·18 $\mathrm{Cdn} = \hat{\Gamma}\{(\exists \Delta) . \Gamma = Q‘\Delta\}$ Df, but the first of these is a temporary definition, for we shall presently identify the class of fields with the class of convex sets.

From 20·21 and 20·11 we have

21·21 $$E'\Gamma \subset C'\Gamma$$

which from ·11 is equivalent to

·22 $$\Gamma \subset F'\Gamma;$$

hence if Γ itself is not null neither is $F'\Gamma$:

·23 $$\exists ! \Gamma = \exists ! F'\Gamma.$$

The existence of $E'\Gamma$ cannot be asserted; for example, if Γ consists of a pair of intersecting straight lines every leaf in the plane contains points of Γ; but we can write conditions for the existence of $E'\Gamma$ in the forms

·24 $$\exists ! E'\Gamma \equiv : (\exists \Upsilon) . \Upsilon \in \text{Clilf} . \Gamma \subset \Upsilon,$$

·25 $$\exists ! E'\Gamma \equiv : (\exists h, \text{K}) . h \in \text{Stl} . \text{K} \in \kappa' C' h . \Gamma \subset \text{K},$$

the second of which merely embodies the definition of a clipped leaf, but permits of a simple translation into words: the field of a set Γ does not occupy the whole plane if there is a straight line which has all the points of Γ on one side of it; what is in fact an equivalent statement is that the field of Γ does not occupy the whole plane if there is a straight line which has no points of Γ on one side of it, which is a translation of

·26 $$\exists ! E'\Gamma \equiv : (\exists \Upsilon) . \Upsilon \in \text{Clilf} . \Gamma \cap \Upsilon = \Lambda,$$

the distinction between this condition and the former being that here we allow points of the set to lie in the bounding line. As with any other boundary

·27 $$\exists ! Q'\Gamma \equiv : \exists ! F'\Gamma . \exists ! E'\Gamma,$$

that is, in virtue of ·23,

·28 $$\exists ! Q'\Gamma \equiv : \exists ! \Gamma . \exists ! E'\Gamma.$$

## 22. *Elementary properties of the field and the cordon.*

For the construction of proofs it is useful to note that from 20·14 and 21·11 the sets $E'\Gamma$, $F'\Gamma$ have the properties

22·11 $$\Gamma \subset \Delta \supset E'\Delta \subset E'\Gamma,$$

·12 $$\Gamma \subset \Delta \supset F'\Gamma \subset F'\Delta.$$

From 21·22 and ·11

·21 $$E'F'\Gamma \subset E'\Gamma;$$

on the other hand, from 20·21 and 21·12

·22 $$\text{M exlf } \Gamma . x \in \text{M} . \supset_x x \in C'F'\Gamma,$$

that is

·23 $$\text{M exlf } \Gamma \supset . F'\Gamma \cap \text{M} = \Lambda,$$

so that

·24 $$\text{M exlf } \Gamma \supset \text{M exlf } F'\Gamma,$$

which is equivalent to

·25 $$(\mathrm{sg}‘\mathrm{exlf})‘\Gamma \subset (\mathrm{sg}‘\mathrm{exlf})‘F‘\Gamma,$$

and implies, from 20·22,

·26 $$E‘\Gamma \subset E‘F‘\Gamma:$$

from ·21, ·26

·27 $$E‘F‘\Gamma = E‘\Gamma,$$

which gives immediately the important result

·28 $$F^2‘\Gamma = F‘\Gamma,$$

shewing that the operator $F$ possesses the property expressed by

·29 $$n \geqslant 1 \supset F^n = F.$$

From the preliminary result 8·33 combined with the definition 20·21 we have

·31 $$x \in y - z \cap E‘\Gamma \supset . \, y \in E‘\Gamma \vee z \in E‘\Gamma,$$

and therefore

·32 $$x \in y - z \, . \, y, z \in F‘\Gamma \, . \supset x \in F‘\Gamma,$$

that is

·33 $$S‘F‘\Gamma \subset F‘\Gamma,$$

or in other terms

·34 $$F‘\Gamma \in \mathrm{Cvx},$$

that is

·35 $$\mathrm{Fld} \subset \mathrm{Cvx};$$

for immediate use we have to note that 21·22, 12·68 and ·33 imply

·36 $$L‘\Gamma \subset F‘\Gamma,$$

and we can sum up propositions ·34 and ·36 in words by saying that whatever the nature of Γ, the field of Γ is a convex set containing all the points and all the primary chords of Γ.

The two propositions ·34 and 15·16 imply

·41 $$F‘\Gamma \in 1 \cup \mathrm{Ds},$$

that is

·42 $$\Gamma \in 1 \, . \vee . \, F‘\Gamma \in \mathrm{Ds}:$$

unless Γ has only one member, $F‘\Gamma$ has no isolated points. Again

·43 $$\mathrm{M} \; \mathrm{exlf} \; \Gamma \, . \, x \in \mathrm{M} \, . \supset x \in D‘\mathrm{M},$$

and therefore

·44 $$x \in E‘\Gamma \supset x \in D‘E‘\Gamma,$$

that is

·45 $$E‘\Gamma \in \mathrm{Ds}:$$

$E‘\Gamma$ has no isolated points. The results ·42, ·45 fall far short of expressing what we really know of the nature of $F\text{‘}\Gamma$ and $E\text{‘}\Gamma$, but they are sufficient grounds for asserting the propositions

22·51 $\Gamma \in 1 \,.\, \mathsf{v} \,.\, G\text{‘}F\text{‘}\Gamma = D\text{‘}F\text{‘}\Gamma,$

·52 $G\text{‘}E\text{‘}\Gamma = D\text{‘}E\text{‘}\Gamma,$

·53 $\Gamma \in 1 \,.\, \mathsf{v} \,.\, Q\text{‘}\Gamma = D\text{‘}F\text{‘}\Gamma \cap D\text{‘}E\text{‘}\Gamma,$

·54 $\Gamma \in 1 \,.\, \mathsf{v} \,.\, Q\text{‘}\Gamma \in \mathrm{Ds},$

·55 $\Gamma \in 1 \,.\, \mathsf{v} \,.\, Q\text{‘}\Gamma \in \mathrm{Pf},$

·56 $\Gamma \in 1 \,.\, \mathsf{v} \,.\, D\text{‘}F\text{‘}\Gamma = F\text{‘}\Gamma \cup Q\text{‘}\Gamma,$

·57 $D\text{‘}E\text{‘}\Gamma = E\text{‘}\Gamma \cup Q\text{‘}\Gamma,$

of which the last two may be put into the forms

·58 $\Gamma \in 1 \,.\, \mathsf{v} \,.\, D\text{‘}F\text{‘}\Gamma = C\text{‘}(E\text{‘}\Gamma - Q\text{‘}\Gamma),$

·59 $D\text{‘}E\text{‘}\Gamma = C\text{‘}(F\text{‘}\Gamma - Q\text{‘}\Gamma).$

## 23. *The domains inside and outside a cordon.*

With reference to a set $\Gamma$ the points of the plane may be divided into four mutually exclusive sets, namely, $F\text{‘}\Gamma - Q\text{‘}\Gamma$, the region of the plane inside the cordon, $F\text{‘}\Gamma \cap Q\text{‘}\Gamma$, the part of the cordon which belongs to the field, $E\text{‘}\Gamma \cap Q\text{‘}\Gamma$, the part of the cordon which does not belong to the field, and $E\text{‘}\Gamma - Q\text{‘}\Gamma$, the region of the plane outside the cordon; there is a close relation between this division of the plane and the division by means of excluding angles into the sets $U\text{‘}\Gamma$, $V\text{‘}\Gamma$, $W\text{‘}\Gamma$. The two sets $F\text{‘}\Gamma - Q\text{‘}\Gamma$ and $E\text{‘}\Gamma - Q\text{‘}\Gamma$ are necessarily domains, and in connection with each of them we use the principle expressed by

23·11 $\Delta \in \mathrm{Dom} \,.\, \Delta \subset \Gamma \,.\, \supset \Delta \subset \Gamma - B\text{‘}\Gamma,$

which gives

·12 $\Delta \in \mathrm{Dom} \,.\, \Delta \subset F\text{‘}\Gamma \,.\, \supset \Delta \subset F\text{‘}\Gamma - Q\text{‘}\Gamma,$

·13 $\Delta \in \mathrm{Dom} \,.\, \Delta \subset E\text{‘}\Gamma \,.\, \supset \Delta \subset E\text{‘}\Gamma - Q\text{‘}\Gamma.$

From 11·22 follows

·21 $U\text{‘}\Gamma \subset \Gamma \cup E\text{‘}\Gamma.$

Since

·22 $x \in V\text{‘}\Gamma \,.\, \Sigma \;\mathrm{les}\,(x, \Gamma) \,.$
$\supset :. (\exists y, z) \,.\, y, z \in \Gamma \,.\, \Sigma \;\mathrm{sect}\,(x \rightarrow y, x \rightarrow z) : \mathsf{v} : (\exists a) \,.\, a \;\mathrm{ef}\; x \,.\, a \;\mathrm{br}\; \Sigma \,.\, \Gamma \cap a = \Lambda,$

we have from 9·64

·23 $x \in V\text{‘}\Gamma$
$\supset :. x \in \Gamma \cup S\text{‘}\Gamma : \mathsf{v} : (\exists a, \Sigma) \,.\, a \;\mathrm{ef}\; x \,.\, a \;\mathrm{br}\; \Sigma \,.\, \mathrm{ang}\text{‘}\Sigma = \pi \,.\, \Gamma \cap (\Sigma \cup a \cup \iota\text{‘}x) = \Lambda,$

so that from 11·23

·24 $V\text{‘}\Gamma \subset L\text{‘}\Gamma \cup E\text{‘}\Gamma;$

moreover, 21·22 and 22·36 shew that in ·21 and ·24 the partition is into mutually exclusive sets. Again, from 11·26

·25 $$x \in E`\Gamma \supset \mathrm{ea}`(x, \Gamma) \geqslant \pi,$$

that is

·26 $$E`\Gamma \subset U`\Gamma \cup V`\Gamma,$$

and this is equivalent to

·27 $$W`\Gamma \subset F`\Gamma.$$

We can now make our applications of ·12 and ·13. From 8·43 we have

·311 $$\Upsilon \in \mathrm{Clilf} \,.\, x \in \Upsilon \,.\, \Gamma \cap \Upsilon = \Lambda \,.\supset x \in E`\Gamma,$$

and so from 8·41 and ·13

·312 $$\Upsilon \in \mathrm{Clilf} \,.\, \Gamma \cap \Upsilon = \Lambda \,.\supset \Upsilon \subset E`\Gamma - Q`\Gamma,$$

which implies

·313 $$\Upsilon \in \mathrm{Clilf} \,.\, x \in B`\Upsilon \,.\, \Gamma \cap \Upsilon = \Lambda \,.\supset x \in D`E`\Gamma,$$

whence, using 11·24 on the one side and 22·57 on the other,

·32 $$\mathrm{ea}`(x, \Gamma) \geqslant \pi \supset x \in E`\Gamma \cup Q`\Gamma,$$

that is

·33 $$U`\Gamma \cup V`\Gamma \subset E`\Gamma \cup Q`\Gamma,$$

which is equivalent to

·34 $$F`\Gamma - Q`\Gamma \subset W`\Gamma.$$

On the other hand, from ·12 with ·27 and 19·46,

·35 $$W`\Gamma \subset F`\Gamma - Q`\Gamma,$$

and this taken with ·34 gives the important result

·36 $$F`\Gamma - Q`\Gamma = W`\Gamma,$$

of which an equivalent form is

·37 $$E`\Gamma \cup Q`\Gamma = U`\Gamma \cup V`\Gamma.$$

We have to notice that ·36 implies

·41 $$F`\Gamma \cup Q`\Gamma = W`\Gamma \cup Q`\Gamma,$$

and that 22·36 and 22·56 give

·42 $$\Gamma \cup D`S`\Gamma \subset F`\Gamma \cup Q`\Gamma,$$

that is

·43 $$\Gamma \cup D`S`\Gamma \subset W`\Gamma \cup Q`\Gamma.$$

From 17·71, 18·35, and 12·37, we have

·51 $$x \in U`\Gamma \cup V`\Gamma \,.\supset :.\, x \in \Gamma \cup D`S`\Gamma \,.\vee :\, (\exists \Upsilon) \,.\, \Upsilon \in \mathrm{Clilf} \,.\, x \in \Upsilon \,.\, \Gamma \cap \Upsilon = \Lambda\,;$$

since $W`\Gamma$ is the complement of $U`\Gamma \cup V`\Gamma$, ·51 and ·43 give

·52 $$x \in U`\Gamma \cup V`\Gamma - Q`\Gamma \supset : (\exists \Upsilon) \,.\, \Upsilon \in \mathrm{Clilf} \,.\, x \in \Upsilon \,.\, \Gamma \cap \Upsilon = \Lambda,$$

which from ·37 is equivalent to

·53 $$E`\Gamma - Q`\Gamma \subset s`\hat{\Upsilon}\{\Upsilon \in \mathrm{Clilf} \,.\, \Gamma \cap \Upsilon = \Lambda\}\,;$$

on the other hand, ·312 is equivalent to

23·54 $$s\text{‘}\hat{\Upsilon}\{\Upsilon \in \text{Clilf} \,.\, \Gamma \cap \Upsilon = \Lambda\} \subset E\text{‘}\Gamma - Q\text{‘}\Gamma,$$

and this combines with the preceding result to give

·55 $$E\text{‘}\Gamma - Q\text{‘}\Gamma = s\text{‘}\hat{\Upsilon}\{\Upsilon \in \text{Clilf} \,.\, \Gamma \cap \Upsilon = \Lambda\}.$$

From 8·42 we have

·56 $$\Upsilon \in \text{Clilf} \,.\, x \in \Upsilon \,.\, \Gamma \cap \Upsilon = \Lambda \,.\supset : (\exists h) \,.\, h \in \text{Stl} \,.\, h \text{ part } (\iota\text{‘}x, \Gamma);$$

since also

·57 $$h \in \text{Stl} \,.\, h \text{ part } (\iota\text{‘}x, \Gamma) \,.\supset : K\text{‘}(x, C\text{‘}h) \in \text{Clilf} \,.\, \Gamma \cap K(x, C\text{‘}h) = \Lambda$$

we have

·58 $$(\exists \Upsilon) \,.\, \Upsilon \in \text{Clilf} \,.\, x \in \Upsilon \,.\, \Gamma \cap \Upsilon = \Lambda : \equiv : (\exists h) \,.\, h \in \text{Stl} \,.\, h \text{ part } (\iota\text{‘}x, \Gamma),$$

and ·55 is equivalent to

·59 $$E\text{‘}\Gamma - Q\text{‘}\Gamma = \hat{x}\{(\exists h) \,.\, h \in \text{Stl} \,.\, h \text{ part } (\iota\text{‘}x, \Gamma)\}.$$

It was in anticipation of ·53 that the preliminary propositions 11·37, 11·38 were proved, for we have from these three results

·61 $$\exists \,!\, U\text{‘}\Gamma \equiv . \, E\text{‘}\Gamma - Q\text{‘}\Gamma = U\text{‘}\Gamma - G\text{‘}\Gamma,$$

·62 $$U\text{‘}\Gamma = \Lambda \,.\, \exists \,!\, E\text{‘}\Gamma \,. \equiv : \Upsilon, \Omega \in \text{Clilf} \,.\, \Gamma \cap \Upsilon = \Lambda \,.\, \Gamma \cap \Omega = \Lambda \,.\supset B\text{‘}\Upsilon \text{ prl } B\text{‘}\Omega.$$

It is easy to prove that the conclusion of ·62 implies that the cordon $Q\text{‘}\Gamma$ is either one straight line or two straight lines: in the first case either $\Gamma$ is contained in this line but not in a ray contained in the line, $F\text{‘}\Gamma$ coincides with $Q\text{‘}\Gamma$, and $E\text{‘}\Gamma - Q\text{‘}\Gamma$ is the complement of $Q\text{‘}\Gamma$, or $F\text{‘}\Gamma - Q\text{‘}\Gamma$ and $E\text{‘}\Gamma - Q\text{‘}\Gamma$ are the parts of the plane lying one on each side of the line; in the second case $F\text{‘}\Gamma - Q\text{‘}\Gamma$ is the strip between the lines and $E\text{‘}\Gamma - Q\text{‘}\Gamma$ is the part of the plane complementary to the sum of the lines and this strip. It is convenient to speak of all these cases and of the case in which $E\text{‘}\Gamma$ does not exist as abnormal, writing

·63 $$\text{Abnlcdn} = \hat{\Gamma}\{(\exists\Delta) \,.\, \Gamma = Q\text{‘}\Delta \,.\, \Gamma = \Lambda \vee \kappa\text{‘}\Gamma \subset \text{Stl}\} \quad \text{Df},$$

·64 $$\text{Nlcdn} = \text{Cdn} - \text{Abnlcdn} \quad \text{Df},$$

definitions not justified until it is shewn that the cordon cannot consist of a number of intersecting lines; then we can write

·65 $$Q\text{‘}\Gamma \in \text{Nlcdn} \equiv : \exists \,!\, E\text{‘}\Gamma \,.\, E\text{‘}\Gamma - Q\text{‘}\Gamma = U\text{‘}\Gamma - G\text{‘}\Gamma.$$

## 24. *Analysis of the cordon of a set.*

There remains the consideration of the points of the cordon itself. One simple expression for the cordon comes directly; from 23·37, 23·51, and 23·55

24·11 $$Q\text{‘}\Gamma \subset (U\text{‘}\Gamma \cup V\text{‘}\Gamma) \cap (\Gamma \cup D\text{‘}S\text{‘}\Gamma),$$

and 23·43 gives the converse inclusion; hence

·12 $$Q\text{‘}\Gamma = (U\text{‘}\Gamma \cup V\text{‘}\Gamma) \cap (\Gamma \cup D\text{‘}S\text{‘}\Gamma),$$

and we may appeal to 12·51 to substitute for this

·13 $$\Gamma \in 1 \,.\vee.\, Q\text{‘}\Gamma = (U\text{‘}\Gamma \cup V\text{‘}\Gamma) \cap D\text{‘}S\text{‘}\Gamma.$$

More detailed results, which follow from 17·71, 18·35, 23·55, and 22·36, are

·14 $$Q‘\Gamma \cap U‘\Gamma \cap F‘\Gamma = U‘\Gamma \cap \Gamma,$$

·15 $$Q‘\Gamma \cap U‘\Gamma \cap E‘\Gamma = U‘\Gamma \cap (D‘\Gamma - \Gamma),$$

·16 $$Q‘\Gamma \cap V‘\Gamma \cap F‘\Gamma = V‘\Gamma \cap L‘\Gamma,$$

·17 $$Q‘\Gamma \cap V‘\Gamma \cap E‘\Gamma = V‘\Gamma \cap (D‘S‘\Gamma - L‘\Gamma);$$

it must not be forgotten that points of all the classes $\Gamma$, $D‘\Gamma - \Gamma$, $L‘\Gamma$, $D‘S‘\Gamma - L‘\Gamma$ may lie inside the cordon although none of these classes have points outside.

From 23·36, 17·23

·21 $$\Gamma \subset \Delta \,.\, \Delta \subset G‘\Gamma \,.\supset F‘\Delta - Q‘\Delta = F‘\Gamma - Q‘\Gamma,$$

and from 23·59, 3·65

·22 $$\Gamma \subset \Delta \,.\, \Delta \subset G‘\Gamma \,.\supset E‘\Delta - Q‘\Delta = E‘\Gamma - Q‘\Gamma,$$

whence

·23 $$\Delta \subset D‘\Gamma \supset Q‘(\Gamma \cup \Delta) = Q‘\Gamma:$$

the cordon of a set is unchanged if the set is enlarged by the addition of any part of its boundary. The cordon of the completed set $G‘\Gamma$ belongs wholly to the field of $G‘\Gamma$, and in terms of this field $Q‘\Gamma$ is given by

·24 $$Q‘\Gamma = (U‘\Gamma \cup V‘\Gamma) \cap F‘G‘\Gamma,$$

while corresponding to ·14, ·16 we have

·25 $$Q‘\Gamma \cap U‘\Gamma = U‘\Gamma \cap G‘\Gamma,$$

·26 $$Q‘\Gamma \cap V‘\Gamma = V‘\Gamma \cap L‘G‘\Gamma.$$

## 25. *The nature of a cordon and of a convex arc.*

Enough has been done to prove that the definitions adopted enable us formally to establish the properties which a cordon obviously possesses. A normal cordon is a complete curve, that is, a united set identical with its own derivative and contained in the derivative of its complement, it has at each point a pair of tangential rays, bounding rays of the limiting excluding sector of the set for that point, which may or may not determine a tangent at the point, and it is a Jordan curve and has a definite finite length between any two of its points. No line that has points of the cordon on both sides cuts the cordon in more than two points; if a line contains a tangential ray of the cordon, the points of the cordon on the line form a stretch, and the remaining points of the cordon are all on the same side of the line.

To constitute a single curve, a set of points must be united. It can however be proved that a set contained in a Jordan curve must be united if it is connected, and therefore a convex arc in a plane may be defined as a connected set contained in its own cordon; such a set need not be complete, for if from a complete convex arc which does not extend to infinity we take an end-point if the arc is not closed or any point whatever if the arc is closed we obtain a convex arc which does not contain one of its limiting points; but a convex arc has at each point which is not an end-point a pair of tangential rays and has between any two of its points a definite finite length.

### 26. *The fundamental properties of the field of a set.*

Returning to the subject of the field, we have from 23·36

26·11 $$F\text{‘}\Gamma \cap (U\text{‘}\Gamma \cup V\text{‘}\Gamma) = F\text{‘}\Gamma \cap Q\text{‘}\Gamma \cap (U\text{‘}\Gamma \cup V\text{‘}\Gamma),$$

and so from 24·14, 24·16, and 17·35

·12 $$F\text{‘}\Gamma \cap (U\text{‘}\Gamma \cup V\text{‘}\Gamma) = L\text{‘}\Gamma \cap (U\text{‘}\Gamma \cup V\text{‘}\Gamma),$$

and since also, from 23·36,

·13 $$F\text{‘}\Gamma \cap W\text{‘}\Gamma = W\text{‘}\Gamma,$$

we have

·14 $$F\text{‘}\Gamma = L\text{‘}\Gamma \cup W\text{‘}\Gamma,$$

which may be expressed, in virtue of 19·45 and 14·22, in the rhetorical form

·15 $$F\text{‘}\Gamma = \Gamma \cup S\text{‘}\Gamma \cup T\text{‘}\Gamma:$$

the points composing the field of a set are the points of the set, the points of its primary chords, and the points of its secondary chords. If we take ·15 with 13·41 we have

·16 $$F\text{‘}\Gamma = L^2\text{‘}\Gamma,$$

or in terms of operators alone

·17 $$F = L^2.$$

The fact that $F\text{‘}\Gamma$ is convex can be written in the form

·21 $$L\text{‘}F\text{‘}\Gamma = F\text{‘}\Gamma,$$

and so by ·17 gives

·22 $$L^3\text{‘}\Gamma = L^2\text{‘}\Gamma,$$

and implies that $L$ has the curious property, possessed also by $B$, the operator giving the boundary of any set, that its complete effect is produced in two operations, though not as a rule in one operation:

·23 $$n \geqslant 2 \supset L^n = L^2,$$

a proposition of which 22·29 is a part: we see now that when we defined primary and secondary chords but not chords of a higher order the limits were not imposed arbitrarily.

In consequence of ·16, we can write 13·58 in the form

·31 $$\Gamma = \Theta \cup \Phi \,.\, \Theta, \Phi \in \mathrm{Cd} \,.\supset F\text{‘}\Gamma = L\text{‘}\Gamma:$$

the field of a set which is connected or is the sum of two connected parts consists of the points belonging to the set and those belonging to its primary chords; the result may be analysed by means of 12·14 and 12·83 into

·32 $$\Gamma \in \mathrm{Cd} \supset .\, F\text{‘}\Gamma = \Gamma \cup S\text{‘}\Gamma,$$

·33 $$\Gamma, \Delta \in \mathrm{Cd} \supset .\, F\text{‘}(\Gamma \cup \Delta) = \Gamma \cup \Delta \cup S\text{‘}\Gamma \cup S\text{‘}\Delta \cup S\text{‘}(\Gamma, \Delta),$$

and we may use the first of these results in the second and write

·34 $$\Gamma, \Delta \in \mathrm{Cd} \supset .\, F\text{‘}(\Gamma \cup \Delta) = F\text{‘}\Gamma \cup F\text{‘}\Delta \cup S\text{‘}(\Gamma, \Delta).$$

It follows from ·15 and 13·31 that

·41 $$F`\Gamma = \Gamma \cup S`\Gamma \cup S^2`\Gamma:$$

also for all values of $n$

·42 $$S^n`\Gamma \subset L^n`\Gamma,$$

and therefore the set $S^n`\Gamma$ is contained in $F`\Gamma$ for all values of $n$, and if $S_*`\Gamma$ denotes the class of sets whose members are $\Gamma$ and all sets of the form $S^n`\Gamma$, the sum $s`S_*`\Gamma$ is contained in $F`\Gamma$; ·41 asserts the converse inclusion, and we have

·43 $$F`\Gamma = s`S_*`\Gamma,$$

a result whose significance will presently be briefly considered.

The value of $F`\Gamma$ in analysis comes largely from a theorem now to be proved, that the field of any set $\Gamma$ is composed of all the points that belong to every convex set containing $\Gamma$; in logical terms, $F`\Gamma$ is the product of the class of convex sets containing $\Gamma$. From 15·12

·51 $$\Gamma \in \mathrm{Cvx} \supset L^2`\Gamma = \Gamma,$$

that is, in virtue of ·16,

·52 $$\Gamma \in \mathrm{Cvx} \supset F`\Gamma = \Gamma:$$

since, by 22·34, $F`\Gamma$ is convex

·53 $$\Gamma = F`\Gamma \supset \Gamma \in \mathrm{Cvx},$$

and combining this result with the preceding

·54 $$\Gamma \in \mathrm{Cvx} \,.\equiv.\, F`\Gamma = \Gamma:$$

a convex set is a set which coincides with its field. Again, from ·52 and 22·12,

·55 $$\Delta \in \mathrm{Cvx} \,.\, \Gamma \subset \Delta \,.\supset F`\Gamma \subset \Delta,$$

which is equivalent to

·56 $$F`\Gamma \subset p`\hat{\Delta}\{\Delta \in \mathrm{Cvx} \,.\, \Gamma \subset \Delta\};$$

also 22·34 and 21·22 together are equivalent to

·57 $$F`\Gamma \in \hat{\Delta}\{\Delta \in \mathrm{Cvx} \,.\, \Gamma \subset \Delta\},$$

which implies

·58 $$p`\hat{\Delta}\{\Delta \in \mathrm{Cvx} \,.\, \Gamma \subset \Delta\} \subset F`\Gamma,$$

and ·56 and ·58 give the desired result,

·59 $$F`\Gamma = p`\hat{\Delta}\{\Delta \in \mathrm{Cvx} \,.\, \Gamma \subset \Delta\},$$

a formula by which $F`\Gamma$ has sometimes been defined.

In order not to interrupt the argument, we did not point out an immediate consequence of ·52 which we announced in 21; from ·52

·61 $$\Gamma \in \mathrm{Cvx} \supset .\, \Delta = \Gamma \supset F`\Delta = \Gamma,$$

and in

·62 $$\Gamma \in \mathrm{Cvx} \supset : (\exists \Delta) \,.\, \Gamma = F`\Delta,$$

a weak deduction from ·61, we have authority for the assertion

26·63 $$\mathrm{Cvx} \subset \mathrm{Fld}$$

which combines with 22·35 to give

·64 $$\mathrm{Fld} = \mathrm{Cvx},$$

whence also

·65 $$\mathrm{Cdn} = B``\mathrm{Cvx};$$

after ·64, use of the contraction Fld is entirely superfluous, but in spite of ·65, Cdn remains of service.

## 27. *Fields, cordons, and convexity with respect to relations in general.*

It is not only for plane sets of points and for the relation denoted in this paper by $S$ that a field is of service, and the merit of 26·43 and 26·59 as definitions of $F`\Gamma$ is that they suggest extensions of the theory of convexity to classes and relations between classes of a very general type. If $R$ is any one-many relation which determines from a class $\alpha$ of any kind another class $R`\alpha$ of the same type as $\alpha$, we can call $\alpha$ convex with respect to the relation $R$ if $R`\alpha$ is contained in $\alpha$, writing

27·11 $$\mathrm{cvx} = \hat{\alpha}\,\hat{R}\,\{R \in 1 \to \mathrm{Cls}\,.\, R`\alpha \subset \alpha\} \quad \mathrm{Df.}$$

If we denote by hyp 27 $a$ the hypothesis that $R$ is such that

hyp 27 $a$ $$\alpha \subset \gamma \supset R`\alpha \subset R`\gamma$$

it can be seen at once that

·12 $$\text{hyp } 27a \;\supset .\; R`\alpha \subset p`\hat{\gamma}\,\{\gamma\,\mathrm{cvx}\,R\,.\,\alpha \subset \gamma\}.$$

Moreover, with the notation already explained in connection with 26·43, $R_*`\alpha$ is a class of classes, and therefore $s`R_*`\alpha$ is a definite class derived from $\alpha$, and $s`R_*$ is like $R$ itself a relation between classes. Again,

·13 $$\text{hyp } 27a \;\supset .\; \gamma\,\mathrm{cvx}\,s`R_* = \gamma\,\mathrm{cvx}\,R,$$

and therefore

·14 $$\text{hyp } 27a \;\supset .\; s`R_*`\alpha \subset p`\hat{\gamma}\,\{\gamma\,\mathrm{cvx}\,R\,.\,\alpha \subset \gamma\}.$$

Whatever the nature of $R$, the class $\alpha$ is a member of the class of classes $R_*`\alpha$ and is contained in $s`R_*`\alpha$, but the assumption hyp 27 $a$ is not sufficient to secure $(s`R_*`\alpha)\,\mathrm{cvx}\,R$, that is

hyp 27 $b$ $$R`s`R_*`\alpha \subset s`R_*`\alpha.$$

We may secure the last condition by a hypothesis *ad hoc*, and enunciate the theorem

·15 $$(\alpha, \gamma)\,.\,\alpha \subset \gamma \supset R`\alpha \subset R`\gamma : (\alpha)\, R`s`R_*`\alpha \subset s`R_*`\alpha :. \supset .\; s`R_*`\alpha = p`\hat{\gamma}\,\{\gamma\,\mathrm{cvx}\,R\,.\,\alpha \subset \gamma\},$$

but if we secure the condition hyp 27 $b$ by the assumption

hyp 27 $c$ $$R`(\alpha \cup \gamma) \subset R`\alpha \cup R`\gamma,$$

and observe that this and the original assumption hyp 27 $a$ together are equivalent to

hyp 27 $d$ $$R`(\alpha \cup \gamma) = R`\alpha \cup R`\gamma,$$

we have the less general but more useful theorem

·16 $(\alpha, \gamma)\, R'(\alpha \cup \gamma) = R'\alpha \cup R'\gamma \,.\supset.\, s'R_*'\alpha = p'\hat{\gamma}\,\{\gamma \operatorname{cvx} R \,.\, \alpha \subset \gamma\}$,

sufficing to shew that with respect to any relation with the property expressed by hyp 27 *d* the field* of a class $\alpha$ may usefully be defined either as a sum or as a product of a class of classes, and the idea of the field may be connected with the idea of convexity with respect to the same relation.

Further, if the classes between which the relation $R$ holds are sets of points in any space in which boundaries exist, we may define the cordon of $\alpha$ with respect to $R$ as $B's'R_*'\alpha$. For example, if $\Gamma$ is a set of points in any reduced† space in which distance is numerical, the field of $\Gamma$ with respect to the operator $D$ that connects $\Gamma$ with its derivative is the completed set $\Gamma \cup D'\Gamma$, and this is the product of the class of complete sets containing $\Gamma$; the cordon $B'G'\Gamma$ of $\Gamma$ with respect to $D$ is not necessarily the same as the boundary $B'\Gamma$ of $\Gamma$, but is in fact composed of those points of $B^2{}'\Gamma$, the boundary of the boundary of $\Gamma$, which are not isolated points of the complement $C'\Gamma$. The utility of a field is certainly not confined to cases in which the relation concerned satisfies hyp 27 *d*, for the relation $S$ defined in 12·11 does not fulfil this condition, and whether the two classes $s'R_*'\alpha$ and $p'\hat{\gamma}\,\{\gamma \operatorname{cvx} R \,.\, \alpha \subset \gamma\}$ are equally important in cases in which they differ, experience alone can decide.

## 28. *Fields in Euclidean space of more than two dimensions.*

Returning more nearly to the main subject of this paper, we observe that in Euclidean space of any finite number of dimensions properties of the field and of the cordon—with respect to the relation of lying in a chord—may be investigated precisely as we have investigated them for a plane, the initial definition of the field being not of either of the general forms 26·43, 26·59 but of a form similar to 21·11, adapted to the geometry of the space and shewn ultimately to be equivalent to a definition in a general form. In three dimensions, the excluding angle is the measure of a dihedral angle, and the points outside the field are defined by means of a standard geometrical figure composed of a leaf together with all the points of space on one side of the plane containing the leaf. It is sufficient to say that in the whole of this theory everything that is obvious is true, but we add that the appearance of the number 2 in 26·16 and 26·23 is associated with the fact that in those propositions plane sets only are in question. In Euclidean space of $m$ dimensions

28·11 $F'\Gamma = L^{m}{}'\Gamma$;

whereas the operator $B$ determining the boundary of a set has the same property§

·12 $B^2 \neq B \,.\, n \geqslant 2 \supset B^n = B^2$

* In work of this general character, the field of a relation has a definite meaning, and although a relation is not a class and there can be no real confusion between the field of a relation and anything which we choose to describe as the field of a class, the use of the word field is open to criticism.

† A space is reduced if of every two points each has a neighbourhood that does not contain the other; in a reduced space with numerical distances the distance between distinct points cannot be zero.

§ Note in formulae ·12, ·13 that to write $R \neq S$ of two one-many relations $R$, $S$ between classes means $(\exists\Gamma) \,.\, R'\Gamma \neq S'\Gamma$, and does not mean $(\Gamma)\, R'\Gamma \neq S'\Gamma$; the relations are not asserted to be mutually exclusive.

in space of all dimensions, the use of Euclidean space of $m$ dimensions makes us acquainted with a simple operator $L$ requiring precisely $m$ applications to secure its ultimate effect, that is, having the property

28·13 $$n \geqslant m \supset L^n = L^m \,.\, n < m \supset L^n \neq L^m.$$

## 29. *Geodesic fields on a sphere.*

A reference to the geometry of a sphere shews that the use of some set analogous to a leaf of a plane is not confined to dealings with the Euclidean plane or with complete Euclidean space of any number of dimensions. If $y$, $z$ are two points of a sphere, let us denote by sa‘$(y, z)$ the points lying between $y$ and $z$ on a great-circle arc subtending an angle not greater than $\pi$ at the centre of the sphere; then writing

$$R‘\Gamma = \hat{x}\{y, z \,\epsilon\, \Gamma \,.\, x \,\epsilon\, \text{sa}‘(y, z)\}$$

we may study the geodesic field of a set $\Gamma$, that is, the field of $\Gamma$ with respect to $R$, and a set $\Gamma$ may be called geodesically convex if $R‘\Gamma$ is contained in $\Gamma$. It is easy to see that a theory analogous to the theory of plane sets developed in this paper may be developed for spherical sets, the place of a leaf being taken by a set which may properly be called a hemisphere, namely, a set composed of all the points on one side of some great circle, together with all the points in a semicircle contained in this great circle, and including one but not both of the end-points of this semicircle.

## 30. *Zeroes of functions of a complex variable: a suggested line of research.*

We conclude with indicating a direction in which research, rendered possible by acquaintance with the concept of the field of a set which is neither finite nor limited, might prove profitable.

It is well known that if $f(z)$ is a polynomial in the complex variable $z$, the points representing the zeroes of the derivative $f'(z)$ belong to the field of the set composed of the zeroes of $f(z)$, and therefore to the field of the zero-set of $f(z) - c$ for any value of $c$. Reference to the functions $\sin z$, $\cos z$ shews that the same result holds for some transcendental functions, and we have only to consider the function $(z-a)/(z-b)^2$ to see that there are functions for which the result does not hold. Denoting the zero-set of the function $f(z)$, that is, the set composed of points in the plane of the complex variable which corresponds to zeroes of the function, by $Z‘\{f(z)\}$, the four questions which first suggest themselves are:

(1) Does the class of functions $f(z)$ such that $Z‘\{f'(z)\}$ is contained in the field of $Z‘\{f(z)\}$ possess any other distinguishing mark?

(2) If $Z‘\{f'(z)\}$ is contained in $F‘Z‘\{f(z)\}$, in what circumstances is it contained in $F‘Z‘\{f(z) - c\}$ for all values of $c$?

(3) If $Z'\{f'(z)\}$ is contained in the field of $Z'\{f(z)-c\}$ for all values of $c$, can there be points outside the field of $Z'\{f'(z)\}$ which also belong to the field of $Z'\{f(z)-c\}$ for all values of $c$?

(4) If $Z'\{f'(z)\}$ is contained in the field of $Z'\{f(z)-c\}$ for some values but not for all values of $c$, does the class of numbers for which the inclusion holds present itself elsewhere in the theory of the function $f(z)$?

There is no difficulty in shewing that for a very large class of integral functions $Z'\{f'(z)\}$ is contained in the field of $Z'\{f(z)-c\}$ for all values of $c$, and that the property is neither common to all integral functions nor peculiar to integral functions, but whether any attempts have been made to answer the third or fourth of the above questions even in the most elementary cases I do not know.

# XIII. *On certain Trigonometrical Sums and their Applications in the Theory of Numbers.*

By S. RAMANUJAN, B.A., F.R.S., Trinity College.

[*Received* and *read* 4 February 1918.]

1. The trigonometrical sums with which this paper is concerned are of the type

$$c_s(n) = \sum_\lambda \cos \frac{2\pi\lambda n}{s},$$

where $\lambda$ is prime to $s$ and not greater than $s$. It is plain that

$$c_s(n) = \Sigma \alpha^n,$$

where $\alpha$ is a primitive root of the equation

$$x^s - 1 = 0.$$

These sums are obviously of very great interest, and a few of their properties have been discussed already*. But, so far as I know, they have never been considered from the point of view which I adopt in this paper; and I believe that all the results which it contains are new.

My principal object is to obtain expressions for a variety of well-known arithmetical functions of $n$ in the form of a series

$$\sum_s a_s c_s(n).$$

A typical formula is

$$\sigma(n) = \frac{\pi^2 n}{6}\left\{\frac{c_1(n)}{1^2} + \frac{c_2(n)}{2^2} + \frac{c_3(n)}{3^2} + \dots\right\},$$

where $\sigma(n)$ is the sum of the divisors of $n$. I give two distinct methods for the proof of this and a large variety of similar formulae. The majority of my formulae are 'elementary' in the technical sense of the word—they can (that is to say) be proved by a combination of processes involving only finite algebra and simple general theorems concerning infinite series. There are however some which are of a 'deeper' character, and can only be proved by means of theorems which seem to depend essentially on the theory of analytic functions. A typical formula of this class is

$$c_1(n) + \tfrac{1}{2}c_2(n) + \tfrac{1}{3}c_3(n) + \dots = 0,$$

a formula which depends upon, and is indeed substantially equivalent to, the 'Prime Number Theorem' of Hadamard and de la Vallée-Poussin.

Many of my formulae are intimately connected with those of my previous paper 'On certain arithmetical functions', published in 1916 in these *Transactions*. They are also connected (in a manner pointed out in § 15) with a joint paper by Mr Hardy and myself, 'Asymptotic Formulae in Combinatory Analysis', in course of publication in the *Proceedings of the London Mathematical Society*.

* See, *e.g.*, Dirichlet-Dedekind, *Vorlesungen über Zahlentheorie*, ed. 4, Supplement VII, pp. 360—370.

2. Let $F(u, v)$ be any function of $u$ and $v$, and let

(2·1) $$D(n) = \sum_{\delta} F(\delta, \delta'),$$

where $\delta$ is a divisor of $n$ and $\delta\delta' = n$. For instance

$$D(1) = F(1, 1); \quad D(2) = F(1, 2) + F(2, 1);$$
$$D(3) = F(1, 3) + F(3, 1); \quad D(4) = F(1, 4) + F(2, 2) + F(4, 1);$$
$$D(5) = F(1, 5) + F(5, 1); \quad D(6) = F(1, 6) + F(2, 3) + F(3, 2) + F(6, 1); \ \ldots\ldots$$

It is clear that $D(n)$ may also be expressed in the form

(2·2) $$D(n) = \sum_{\delta} F(\delta', \delta).$$

Suppose now that

(2·3) $$\eta_s(n) = \sum_{0}^{s-1} \cos \frac{2\pi\nu n}{s},$$

so that $\eta_s(n) = s$ if $s$ is a divisor of $n$ and $\eta_s(n) = 0$ otherwise. Then

(2·4) $$D(n) = \sum_{1}^{t} \frac{1}{\nu} \eta_\nu(n) F\left(\nu, \frac{n}{\nu}\right)^*,$$

where $t$ is any number not less than $n$. Now let

(2·5) $$c_s(n) = \sum_{\lambda} \cos \frac{2\pi\lambda n}{s},$$

where $\lambda$ is prime to $s$ and does not exceed $s$; *e.g.*

$$c_1(n) = 1; \ c_2(n) = \cos n\pi; \ c_3(n) = 2\cos \tfrac{2}{3}n\pi;$$
$$c_4(n) = 2\cos \tfrac{1}{2}n\pi; \ c_5(n) = 2\cos \tfrac{2}{5}n\pi + 2\cos \tfrac{4}{5}n\pi;$$
$$c_6(n) = 2\cos \tfrac{1}{3}n\pi; \ c_7(n) = 2\cos \tfrac{2}{7}n\pi + 2\cos \tfrac{4}{7}n\pi + 2\cos \tfrac{6}{7}n\pi;$$
$$c_8(n) = 2\cos \tfrac{1}{4}n\pi + 2\cos \tfrac{3}{4}n\pi; \ c_9(n) = 2\cos \tfrac{2}{9}n\pi + 2\cos \tfrac{4}{9}n\pi + 2\cos \tfrac{8}{9}n\pi;$$
$$c_{10}(n) = 2\cos \tfrac{1}{5}n\pi + 2\cos \tfrac{3}{5}n\pi; \ \ldots\ldots$$

It follows from (2·3) and (2·5) that

(2·6) $$\eta_s(n) = \sum_{\delta} c_\delta(n),$$

where $\delta$ is a divisor of $s$; and hence† that

(2·7) $$c_s(n) = \sum_{\delta} \mu(\delta') \eta_\delta(n),$$

where $\delta$ is a divisor of $s$, $\delta\delta' = s$, and

(2·8) $$\sum \frac{\mu(\nu)}{\nu^s} = \frac{1}{\zeta(s)},$$

$\zeta(s)$ being the Riemann Zeta-function. In particular

$$c_1(n) = \eta_1(n); \ c_2(n) = \eta_2(n) - \eta_1(n); \ c_3(n) = \eta_3(n) - \eta_1(n);$$
$$c_4(n) = \eta_4(n) - \eta_2(n); \ c_5(n) = \eta_5(n) - \eta_1(n); \ \ldots\ldots$$

* $\sum_{1}^{t}$ is to be understood as meaning $\sum_{1}^{[t]}$, where $[t]$ denotes as usual the greatest integer in $t$.

† See Landau, *Handbuch der Lehre von der Verteilung der Primzahlen*, p. 577.

But from (2·3) we know that $\eta_\delta(n) = 0$ if $\delta$ is not a divisor of $n$; and so we can suppose that, in (2·7), $\delta$ is a common divisor of $n$ and $s$. It follows that

$$c_s(n) \leqslant \Sigma\delta,$$

where $\delta$ is a divisor of $n$; so that

(2·9) $$c_\nu(n) = O(1)$$

*if $n$ is fixed and $\nu \to \infty$.* Since

$$\eta_s(n) = \eta_s(n+s); \quad c_s(n) = c_s(n+s),$$

the values of $c_s(n)$ for $n = 1, 2, 3 \ldots$ can be shown conveniently by writing

$$c_1(n) = \overline{1}; \quad c_2(n) = \overline{-1, 1}; \quad c_3(n) = \overline{-1, -1, 2};$$

$$c_4(n) = \overline{0, -2, 0, 2}; \quad c_5(n) = \overline{-1, -1, -1, -1, 4};$$

$$c_6(n) = \overline{1, -1, -2, -1, 1, 2}; \quad c_7(n) = \overline{-1, -1, -1, -1, -1, -1, 6};$$

$$c_8(n) = \overline{0, 0, 0, -4, 0, 0, 0, 4}; \quad c_9(n) = \overline{0, 0, -3, 0, 0, -3, 0, 0, 6};$$

$$c_{10}(n) = \overline{1, -1, 1, -1, -4, -1, 1, -1, 1, 4}; \quad \ldots\ldots$$

the meaning of the third formula, for example, being that $c_3(1) = -1$, $c_3(2) = -1$, $c_3(3) = 2$, and that these values are then repeated periodically.

It is plain that we have also

(2·91) $$c_\nu(n) = O(1),$$

*when $\nu$ is fixed and $n \to \infty$.*

3. Substituting (2·6) in (2·4), and collecting the coefficients of $c_1(n)$, $c_2(n)$, $c_3(n)$, ..., we find that

(3·1) $$D(n) = c_1(n)\sum_1^t \frac{1}{\nu} F\left(\nu, \frac{n}{\nu}\right) + c_2(n) \sum_1^{\frac{1}{2}t} \frac{1}{2\nu} F\left(2\nu, \frac{n}{2\nu}\right) + c_3(n) \sum_1^{\frac{1}{3}t} \frac{1}{3\nu} F\left(3\nu, \frac{n}{3\nu}\right) + \ldots,$$

where $t$ is any number not less than $n$. If we use (2·2) instead of (2·1) we obtain another expression, viz.

(3·2) $$D(n) = c_1(n)\sum_1^t \frac{1}{\nu} F\left(\frac{n}{\nu}, \nu\right) + c_2(n) \sum_1^{\frac{1}{2}t} \frac{1}{2\nu} F\left(\frac{n}{2\nu}, 2\nu\right) + c_3(n) \sum_1^{\frac{1}{3}t} \frac{1}{3\nu} F\left(\frac{n}{3\nu}, 3\nu\right) + \ldots,$$

where $t$ is any number not less than $n$.

Suppose now that

$$F_1(u, v) = F(u, v) \log u, \quad F_2(u, v) = F(u, v) \log v.$$

Then we have

$$D(n) \log n = \sum_\delta F(\delta, \delta') \log n = \sum_\delta F(\delta, \delta') \log(\delta\delta')$$

$$= \sum_\delta F_1(\delta, \delta') + \sum_\delta F_2(\delta, \delta'),$$

where $\delta$ is a divisor of $n$ and $\delta\delta' = n$.

Now for $\sum_\delta F_1(\delta, \delta')$ we shall write the expression corresponding to (3·1) and for $\sum_\delta F_2(\delta, \delta')$ the expression corresponding to (3·2). Then we have

(3·3) $$D(n)\log n = c_1(n)\sum_1^r \frac{\log \nu}{\nu} F\left(\nu, \frac{n}{\nu}\right) + c_2(n)\sum_1^{\frac{1}{2}r} \frac{\log 2\nu}{2\nu} F\left(2\nu, \frac{n}{2\nu}\right) + c_3(n)\sum_1^{\frac{1}{3}r} \frac{\log 3\nu}{3\nu} F\left(3\nu, \frac{n}{3\nu}\right) + \ldots$$
$$+ c_1(n)\sum_1^t \frac{\log \nu}{\nu} F\left(\frac{n}{\nu}, \nu\right) + c_2(n)\sum_1^{\frac{1}{2}t} \frac{\log 2\nu}{2\nu} F\left(\frac{n}{2\nu}, 2\nu\right) + c_3(n)\sum_1^{\frac{1}{3}t} \frac{\log 3\nu}{3\nu} F\left(\frac{n}{3\nu}, 3\nu\right) + \ldots,$$

where $r$ and $t$ are any two numbers not less than $n$. If, in particular, $F(u, v) = F(v, u)$, then (3·3) reduces to

(3·4) $$\tfrac{1}{2} D(n)\log n = c_1(n)\sum_1^t \frac{\log \nu}{\nu} F\left(\nu, \frac{n}{\nu}\right) + c_2(n)\sum_1^{\frac{1}{2}t} \frac{\log 2\nu}{2\nu} F\left(2\nu, \frac{n}{2\nu}\right) + c_3(n)\sum_1^{\frac{1}{3}t} \frac{\log 3\nu}{3\nu} F\left(3\nu, \frac{n}{3\nu}\right) + \ldots,$$

where $t$ is any number not less than $n$.

4. We may also write $D(n)$ in the form

(4·1) $$D(n) = \sum_{\delta=1}^{u} F(\delta, \delta') + \sum_{\delta=1}^{v} F(\delta', \delta),$$

where $\delta$ is a divisor of $n$, $\delta\delta' = n$, and $u$, $v$ are any two positive numbers such that $uv = n$, it being understood that, if $u$ and $v$ are both integral, a term $F(u, v)$ is to be subtracted from the right-hand side. Hence (with the same conventions)

$$D(n) = \sum_1^u \frac{1}{\nu}\eta_\nu(n) F\left(\nu, \frac{n}{\nu}\right) + \sum_1^v \frac{1}{\nu}\eta_\nu(n) F\left(\frac{n}{\nu}, \nu\right).$$

Applying to this formula transformations similar to those of § 3, we obtain

(4·2) $$D(n) = c_1(n)\sum_1^u \frac{1}{\nu} F\left(\nu, \frac{n}{\nu}\right) + c_2(n)\sum_1^{\frac{1}{2}u} \frac{1}{2\nu} F\left(2\nu, \frac{n}{2\nu}\right) + \ldots$$
$$+ c_1(n)\sum_1^v \frac{1}{\nu} F\left(\frac{n}{\nu}, \nu\right) + c_2(n)\sum_1^{\frac{1}{2}v} \frac{1}{2\nu} F\left(\frac{n}{2\nu}, 2\nu\right) + \ldots.$$

where $u$ and $v$ are positive numbers such that $uv = n$. If $u$ and $v$ are integers then a term $F(u, v)$ should be subtracted from the right-hand side.

If we suppose that $0 < u \leqslant 1$ then (4·2) reduces to (3·2), and if $0 < v \leqslant 1$ it reduces to (3·1). Another particular case of interest is that in which $u = v$. Then

(4·3) $$D(n) = c_1(n)\sum_1^{\sqrt{n}} \frac{1}{\nu}\left\{F\left(\nu, \frac{n}{\nu}\right) + F\left(\frac{n}{\nu}, \nu\right)\right\} + c_2(n)\sum_1^{\frac{1}{2}\sqrt{n}} \frac{1}{2\nu}\left\{F\left(2\nu, \frac{n}{2\nu}\right) + F\left(\frac{n}{2\nu}, 2\nu\right)\right\} + \ldots.$$

If $n$ is a perfect square then $F(\sqrt{n}, \sqrt{n})$ should be subtracted from the right-hand side.

5. We shall now consider some special forms of these general equations. Suppose that $F(u, v) = v^s$, so that $D(n)$ is the sum $\sigma_s(n)$ of the $s$th powers of the divisors of $n$. Then from (3·1) and (3·2) we have

(5·1) $$\frac{\sigma_s(n)}{n^s} = c_1(n)\sum_1^t \frac{1}{\nu^{s+1}} + c_2(n)\sum_1^{\frac{1}{2}t} \frac{1}{(2\nu)^{s+1}} + c_3(n)\sum_1^{\frac{1}{3}t} \frac{1}{(3\nu)^{s+1}} + \ldots$$

(5·2) $$\sigma_s(n) = c_1(n)\sum_1^t \nu^{s-1} + c_2(n)\sum_1^{\frac{1}{2}t} (2\nu)^{s-1} + c_3(n)\sum_1^{\frac{1}{3}t} (3\nu)^{s-1} + \ldots.$$

where $t$ is any number not less than $n$: from (3·3)

$$(5\cdot3) \qquad \sigma_s(n)\log n = c_1(n)\sum_1^r \nu^{s-1}\log\nu + c_2(n)\sum_1^{\frac{1}{2}r}(2\nu)^{s-1}\log 2\nu + \dots$$
$$+ n^s\left\{c_1(n)\sum_1^t \frac{\log\nu}{\nu^{s+1}} + c_2(n)\sum_1^{\frac{1}{2}t}\frac{\log 2\nu}{(2\nu)^{s+1}} + \dots\right\},$$

where $r$ and $t$ are any two numbers not less than $n$: and from (4·2)

$$(5\cdot4) \qquad \sigma_s(n) = c_1(n)\sum_1^u \nu^{s-1} + c_2(n)\sum_1^{\frac{1}{2}u}(2\nu)^{s-1} + c_3(n)\sum_1^{\frac{1}{3}u}(3\nu)^{s-1} + \dots$$
$$+ n^s\left\{c_1(n)\sum_1^v \frac{1}{\nu^{s+1}} + c_2(n)\sum_1^{\frac{1}{2}v}\frac{1}{(2\nu)^{s+1}} + c_3(n)\sum_1^{\frac{1}{3}v}\frac{1}{(3\nu)^{s+1}} + \dots\right\},$$

where $uv = n$. If $u$ and $v$ are integers then $u^s$ should be subtracted from the right-hand side.

Let $d(n) = \sigma_0(n)$ denote the number of divisors of $n$ and $\sigma(n) = \sigma_1(n)$ the sum of the divisors of $n$. Then from (5·1)—(5·4) we obtain

$$(5\cdot5) \qquad d(n) = c_1(n)\sum_1^t \frac{1}{\nu} + c_2(n)\sum_1^{\frac{1}{2}t}\frac{1}{2\nu} + c_3(n)\sum_1^{\frac{1}{3}t}\frac{1}{3\nu} + \dots,$$

$$(5\cdot6) \qquad \sigma(n) = c_1(n)[t] + c_2(n)[\tfrac{1}{2}t] + c_3(n)[\tfrac{1}{3}t] + \dots.$$

$$(5\cdot7) \qquad \tfrac{1}{2}d(n)\log n = c_1(n)\sum_1^t \frac{\log\nu}{\nu} + c_2(n)\sum_1^{\frac{1}{2}t}\frac{\log 2\nu}{2\nu} + c_3(n)\sum_1^{\frac{1}{3}t}\frac{\log 3\nu}{3\nu} + \dots,$$

$$(5\cdot8) \qquad d(n) = c_1(n)\left\{\sum_1^u \frac{1}{\nu} + \sum_1^v \frac{1}{\nu}\right\} + c_2(n)\left\{\sum_1^{\frac{1}{2}u}\frac{1}{2\nu} + \sum_1^{\frac{1}{2}v}\frac{1}{2\nu}\right\} + c_3(n)\left\{\sum_1^{\frac{1}{3}u}\frac{1}{3\nu} + \sum_1^{\frac{1}{3}v}\frac{1}{3\nu}\right\} + \dots.$$

where $t \geqslant n$ and $uv = n$. If $u$ and $v$ are integers then 1 should be subtracted from the right-hand side of (5·8). Putting $u = v = \sqrt{n}$ in (5·8) we obtain

$$(5\cdot9) \qquad \tfrac{1}{2}d(n) = c_1(n)\sum_1^{\sqrt{n}}\frac{1}{\nu} + c_2(n)\sum_1^{\frac{1}{2}\sqrt{n}}\frac{1}{2\nu} + c_3(n)\sum_1^{\frac{1}{3}\sqrt{n}}\frac{1}{3\nu} + \dots,$$

unless $n$ is a perfect square, when $\frac{1}{2}$ should be subtracted from the right-hand side. It may be interesting to note that, if we replace the left-hand side in (5·9) by

$$[\tfrac{1}{2} + \tfrac{1}{2}d(n)],$$

then the formula is true without exception.

6. So far our work has been based on elementary formal transformations, and no questions of convergence have arisen. We shall now consider the equation (5·1) more carefully. Let us suppose that $s > 0$. Then

$$\sum_1^{t/k}\frac{1}{(k\nu)^{s+1}} = \sum_1^{\infty}\frac{1}{(k\nu)^{s+1}} + O\left(\frac{1}{kt^s}\right) = \frac{1}{k^{s+1}}\zeta(s+1) + O\left(\frac{1}{kt^s}\right).$$

The number of terms in the right-hand side of (5·1) is $[t]$. Also we know that $c_\nu(n) = O(1)$ as $\nu \to \infty$. Hence

$$\frac{\sigma_s(n)}{n^s} = \zeta(s+1)\sum_{\nu=1}^t \frac{c_\nu(n)}{\nu^{s+1}} + O\left\{\frac{1}{t^s}\sum_{\nu=1}^t \frac{1}{\nu}\right\}$$
$$= \zeta(s+1)\sum_1^{\infty}\frac{c_\nu(n)}{\nu^{s+1}} + O\left(\frac{\log t}{t^s}\right).$$

Making $t \to \infty$ we obtain

$$(6\cdot1) \qquad \sigma_s(n) = n^s \zeta(s+1) \left\{ \frac{c_1(n)}{1^{s+1}} + \frac{c_2(n)}{2^{s+1}} + \frac{c_3(n)}{3^{s+1}} + \dots \right\},$$

if $s > 0$. Similarly, if we make $t \to \infty$ in (5·3), we obtain

$$\sigma_s(n) \log n = c_1(n) \sum_1^t \nu^{s-1} \log \nu + c_2(n) \sum_1^{\frac{1}{2}t} (2\nu)^{s-1} \log 2\nu + \dots$$
$$+ n^s \left\{ c_1(n) \sum_1^\infty \frac{\log \nu}{\nu^{s+1}} + c_2(n) \sum_1^\infty \frac{\log 2\nu}{(2\nu)^{s+1}} + \dots \right\}.$$

But

$$\sum_1^\infty \frac{\log k\nu}{(k\nu)^{s+1}} = \frac{\log k}{k^{s+1}} \zeta(s+1) - \frac{1}{k^{s+1}} \zeta'(s+1).$$

It follows from this and (6·1) that

$$(6\cdot2) \qquad \sigma_s(n) \left\{ \frac{\zeta'(s+1)}{\zeta(s+1)} + \log n \right\} = c_1(n) \sum_1^t \nu^{s-1} \log \nu + c_2(n) \sum_1^{\frac{1}{2}t} (2\nu)^{s-1} \log 2\nu + \dots$$
$$+ n^s \zeta(s+1) \left\{ \frac{c_1(n) \log 1}{1^{s+1}} + \frac{c_2(n) \log 2}{2^{s+1}} + \frac{c_3(n) \log 3}{3^{s+1}} + \dots \right\},$$

where $s > 0$ and $t \geqslant n$. Putting $s = 1$ in (6·1) and (6·2) we obtain

$$(6\cdot3) \qquad \sigma(n) = \frac{\pi^2}{6} n \left\{ \frac{c_1(n)}{1^2} + \frac{c_2(n)}{2^2} + \frac{c_3(n)}{3^2} + \dots \right\},$$

$$(6\cdot4) \qquad \sigma(n) \left\{ \frac{\zeta'(2)}{\zeta(2)} + \log n \right\} = \frac{\pi^2}{6} n \left\{ \frac{c_1(n)}{1^2} \log 1 + \frac{c_2(n)}{2^2} \log 2 + \dots \right\}$$
$$+ c_1(n) [t] \log 1 + c_2(n) [\tfrac{1}{2}t] \log 2 + \dots$$
$$+ c_1(n) \log [t]! + c_2(n) \log [\tfrac{1}{2}t]! + \dots,$$

where $t \geqslant n$.

7. Since

$$(7\cdot1) \qquad \sigma_s(n) = n^s \sigma_{-s}(n),$$

we may write (6·1) in the form

$$(7\cdot2) \qquad \frac{\sigma_{-s}(n)}{\zeta(s+1)} = \frac{c_1(n)}{1^{s+1}} + \frac{c_2(n)}{2^{s+1}} + \frac{c_3(n)}{3^{s+1}} + \dots,$$

where $s > 0$. This result has been proved by purely elementary methods. But in order to know whether the right-hand side of (7·2) is convergent or not for values of $s$ less than or equal to zero we require the help of theorems which have only been established by transcendental methods.

Now the right-hand side of (7·2) is an ordinary Dirichlet's series for

$$\sigma_{-s}(n) \times \frac{1}{\zeta(s+1)}.$$

The first factor is a finite Dirichlet's series and so an absolutely convergent Dirichlet's series.

It follows that the right-hand side of (7·2) is convergent whenever the Dirichlet's series for $1/\zeta(s+1)$, viz.

$$\Sigma \frac{\mu(n)}{n^{1+s}}, \tag{7·3}$$

is convergent. But it is known* that the series (7·3) is convergent when $s=0$ and that its sum is 0.

It follows from this that

$$c_1(n) + \tfrac{1}{2}c_2(n) + \tfrac{1}{3}c_3(n) + \ldots = 0. \tag{7·4}$$

Nothing is known about the convergence of (7·3) when $-\frac{1}{2} < s < 0$. But with the assumption of the truth of the hitherto unproved Riemann hypothesis it has been proved† that (7·3) is convergent when $s > -\frac{1}{2}$. With this assumption we see that (7·2) is true when $s > -\frac{1}{2}$. In other words, if $-\frac{1}{2} < s < \frac{1}{2}$ then

$$\begin{aligned} \sigma_s(n) &= \zeta(1-s)\left\{\frac{c_1(n)}{1^{1-s}} + \frac{c_2(n)}{2^{1-s}} + \frac{c_3(n)}{3^{1-s}} + \ldots\right\} \\ &= n^s\zeta(1+s)\left\{\frac{c_1(n)}{1^{1+s}} + \frac{c_2(n)}{2^{1-s}} + \frac{c_3(n)}{3^{1-s}} + \ldots\right\}. \end{aligned} \tag{7·5}$$

8. It is known‡ that all the series obtained from (7·3) by term-by-term differentiation with respect to $s$ are convergent when $s=0$; and it is obvious that the derivatives of $\sigma_{-s}(n)$ with respect to $s$ are all finite Dirichlet's series and so absolutely convergent. It follows that all the derivatives of the right-hand side of (7·2) are convergent when $s=0$; and so we can equate the coefficients of like powers of $s$ from the two sides of (7·2). Now

$$\frac{1}{\zeta(s+1)} = s - \gamma s^2 + \ldots, \tag{8·1}$$

where $\gamma$ is Euler's constant. And

$$\sigma_{-s}(n) = \sum_{\delta} \delta^{-s} = \sum_{\delta} 1 - s\sum_{\delta} \log \delta + \ldots,$$

where $\delta$ is a divisor of $n$. But

$$\sum_{\delta} \log \delta = \sum_{\delta} \log \delta' = \tfrac{1}{2}\sum_{\delta} \log(\delta\delta') = \tfrac{1}{2}d(n)\log n,$$

where $\delta\delta' = n$. Hence

$$\sigma_{-s}(n) = d(n) - \tfrac{1}{2}sd(n)\log n + \ldots. \tag{8·2}$$

Now equating the coefficients of $s$ and $s^2$ from the two sides of (7·2), and using (8·1) and (8·2), we obtain

$$c_1(n)\log 1 + \tfrac{1}{2}c_2(n)\log 2 + \tfrac{1}{3}c_3(n)\log 3 + \ldots = -d(n), \tag{8·3}$$

$$c_1(n)(\log 1)^2 + \tfrac{1}{2}c_2(n)(\log 2)^2 + \tfrac{1}{3}c_3(n)(\log 3)^2 + \ldots = -d(n)(2\gamma + \log n). \tag{8·4}$$

9. I shall now find an expression of the same kind for $\phi(n)$, the number of numbers prime to and not exceeding $n$. Let $p_1, p_2, p_3, \ldots$ be the prime divisors of $n$ and let

$$\phi_s(n) = n^s(1-p_1^{-s})(1-p_2^{-s})(1-p_3^{-s})\ldots, \tag{9·1}$$

so that $\phi_1(n) = \phi(n)$. Suppose that

$$F(u, v) = \mu(u)v^s.$$

* Landau, *Handbuch*, p. 591.

† Littlewood, *Comptes Rendus*, 29 Jan. 1912.

‡ Landau, *Handbuch*, p. 594.

Then it is easy to see that

$$D(n) = \phi_s(n).$$

Hence, from (3·1), we have

$$\frac{\phi_s(n)}{n^s} = c_1(n) \sum_1^t \frac{\mu(\nu)}{\nu^{s+1}} + c_2(n) \sum_1^{\frac{1}{2}t} \frac{\mu(2\nu)}{(2\nu)^{s+1}} + \dots, \tag{9·2}$$

where $t$ is any number not less than $n$. If $s > 0$ we can make $t \to \infty$, as in § 6. Then we have

$$\frac{\phi_s(n)}{n^s} = c_1(n) \sum_1^\infty \frac{\mu(\nu)}{\nu^{s+1}} + c_2(n) \sum_1^\infty \frac{\mu(2\nu)}{(2\nu)^{s+1}} + \dots. \tag{9·3}$$

But it can easily be shown that

$$\sum_1^\infty \frac{\mu(n\nu)}{\nu^s} = \frac{\mu(n)}{\zeta(s)(1-p_1^{-s})(1-p_2^{-s})(1-p_3^{-s})\dots}, \tag{9·4}$$

where $p_1, p_2, p_3, \dots$ are the prime divisors of $n$. In other words

$$\sum_1^\infty \frac{\mu(n\nu)}{\nu^s} = \frac{\mu(n)\, n^s}{\phi_s(n)\, \zeta(s)}. \tag{9·5}$$

It follows from (9·3) and (9·5) that

$$\frac{\phi_s(n)\,\zeta(s+1)}{n^s} = \frac{\mu(1)\,c_1(n)}{\phi_{s+1}(1)} + \frac{\mu(2)\,c_2(n)}{\phi_{s+1}(2)} + \frac{\mu(3)\,c_3(n)}{\phi_{s+1}(3)} + \dots. \tag{9·6}$$

In particular

$$\frac{\pi^2}{6n}\phi(n) = c_1(n) - \frac{c_2(n)}{2^2-1} - \frac{c_3(n)}{3^2-1} - \frac{c_5(n)}{5^2-1} + \frac{c_6(n)}{(2^2-1)(3^2-1)} - \frac{c_7(n)}{7^2-1} + \frac{c_{10}(n)}{(2^2-1)(5^2-1)} - \dots. \tag{9·7}$$

10. I shall now consider an application of the main formulae to the problem of the number of representations of a number as the sum of 2, 4, 6, 8, ... squares. We shall require the following preliminary results.

(1) Let

$$\Sigma D(n)\, x^n = X_1 = \frac{1^{s-1}x}{1+x} + \frac{2^{s-1}x^2}{1-x^2} + \frac{3^{s-1}x^3}{1+x^3} + \dots. \tag{10·1}$$

We shall choose

$$\begin{aligned} F(u, v) &= v^{s-1}, & u &\equiv 1 \pmod 2),\\ F(u, v) &= -v^{s-1}, & u &\equiv 2 \pmod 4),\\ F(u, v) &= (2^s-1)\, v^{s-1}, & u &\equiv 0 \pmod 4). \end{aligned}$$

Then from (3·1) we can show, by arguments similar to those used in § 6, that

$$D(n) = n^{s-1}(1^{-s}+3^{-s}+5^{-s}+\dots)\{1^{-s}c_1(n)+2^{-s}c_4(n)+3^{-s}c_3(n)+4^{-s}c_8(n) + 5^{-s}c_5(n)+6^{-s}c_{12}(n)+7^{-s}c_7(n)+8^{-s}c_{16}(n)+\dots\} \tag{10·11}$$

if $s > 1$.

(2) Let

$$\Sigma D(n)\, x^n = X_2 = \frac{1^{s-1}x}{1-x} + \frac{2^{s-1}x^2}{1+x^2} + \frac{3^{s-1}x^3}{1-x^3} + \dots. \tag{10·2}$$

We shall choose

$$F(u, v) = v^{s-1}, \qquad u \equiv 1 \pmod 2,$$
$$F(u, v) = v^{s-1}, \qquad u \equiv 2 \pmod 4,$$
$$F(u, v) = (1 - 2^s)\, v^{s-1}, \quad u \equiv 0 \pmod 4.$$

Then we obtain as before

(10·21) $$D(n) = n^{s-1}(1^{-s} + 3^{-s} + 5^{-s} + \ldots)\{1^{-s} c_1(n) - 2^{-s} c_4(n) + 3^{-s} c_3(n) - 4^{-s} c_8(n) + 5^{-s} c_5(n) - 6^{-s} c_{12}(n) + 7^{-s} c_7(n) - 8^{-s} c_{16}(n) + \ldots\}.$$

(3) Let

(10·3) $$\Sigma D(n) x^n = X_3 = \frac{1^{s-1} x}{1 + x^2} + \frac{2^{s-1} x^2}{1 + x^4} + \frac{3^{s-1} x^3}{1 + x^6} + \ldots.$$

We shall choose

$$F(u, v) = 0, \qquad u \equiv 0 \pmod 2,$$
$$F(u, v) = v^{s-1}, \qquad u \equiv 1 \pmod 4,$$
$$F(u, v) = -v^{s-1}, \quad u \equiv 3 \pmod 4.$$

Then we obtain as before

(10·31) $$D(n) = n^{s-1}(1^{-s} - 3^{-s} + 5^{-s} - \ldots)\{1^{-s} c_1(n) - 3^{-s} c_3(n) + 5^{-s} c_5(n) - \ldots\}.$$

(4) We shall also require a similar formula for the function $D(n)$ defined by

(10·4) $$\Sigma D(n) x^n = X_4 = \frac{1^{s-1} x}{1 - x} - \frac{3^{s-1} x^3}{1 - x^3} + \frac{5^{s-1} x^5}{1 - x^5} - \ldots.$$

The formula required is not a direct consequence of the preceding analysis, but if, instead of starting with the function

$$c_r(n) = \sum_{\lambda} \cos \frac{2\pi n\lambda}{r},$$

we start with the function

$$s_r(n) = \sum_{\lambda} (-1)^{\frac{1}{2}(\lambda - 1)} \sin \frac{2\pi n\lambda}{r},$$

where $\lambda$ is prime to $r$ and does not exceed $r$, and proceed as in §§ 2–3, we can show that

(10·41) $$D(n) = \tfrac{1}{2} n^{s-1}(1^{-s} - 3^{-s} + 5^{-s} - \ldots)\{1^{-s} s_4(n) + 2^{-s} s_8(n) + 3^{-s} s_{12}(n) + \ldots\}.$$

It should be observed that there is a correspondence between $c_r(n)$ and the ordinary $\zeta$-function on the one hand and $s_r(n)$ and the function

$$\eta(s) = 1^{-s} - 3^{-s} + 5^{-s} - \ldots$$

on the other. It is possible to define an infinity of systems of trigonometrical sums such as $c_r(n)$, $s_r(n)$, each corresponding to one of the general class of '$L$-functions*' of which $\zeta(s)$ and $\eta(s)$ are the simplest members.

We have shown that (10·31) and (10·41) are true when $s > 1$. But if we assume that the Dirichlet's series for $1/\eta(s)$ is convergent when $s = 1$, a result which is precisely of the same depth as the prime number theorem and has only been established by transcendental methods, then we can show by arguments similar to those of § 7 that (10·31) and (10·41) are true when $s = 1$.

* See Landau, *Handbuch*, pp. 414 *et seq.*

11. I have shown elsewhere* that if $s$ is a positive integer and

$$1 + \Sigma\, r_s(n)\, x^n = (1 + 2x + 2x^4 + 2x^9 + \dots)^s,$$

then

$$r_{2s}(n) = \delta_{2s}(n) + e_{2s}(n),$$

where $e_{2s}(n) = 0$ when $s = 1$, 2, 3 or 4 and is of lower order† than $\delta_{2s}(n)$ in all cases; that if $s$ is a multiple of 4 then

$$(11{\cdot}1) \qquad (1^{-s} + 3^{-s} + 5^{-s} + \dots)\, \Sigma\, \delta_{2s}(n)\, x^n = \frac{\pi^s}{(s-1)!} X_1;$$

if $s$ is of the form $4k + 2$ then

$$(11{\cdot}2) \qquad (1^{-s} + 3^{-s} + 5^{-s} + \dots)\, \Sigma\, \delta_{2s}(n)\, x^n = \frac{\pi^s}{(s-1)!} X_2;$$

if $s$ is of the form $4k + 1$ then

$$(11{\cdot}3) \qquad (1^{-s} - 3^{-s} + 5^{-s} - \dots)\, \Sigma\, \delta_{2s}(n)\, x^n = \frac{\pi^s}{(s-1)!} (X_3 + 2^{1-s} X_4),$$

except when $s = 1$; and if $t$ is of the form $4k + 3$ then

$$(11{\cdot}4) \qquad (1^{-s} - 3^{-s} + 5^{-s} - \dots)\, \Sigma\, \delta_{2s}(n)\, x^n = \frac{\pi^s}{(s-1)!} (X_3 - 2^{1-s} X_4),$$

$X_1$, $X_2$, $X_3$, $X_4$ being the same as in § 10.

In the case in which $s = 1$ it is well known that

$$(11{\cdot}5) \qquad \Sigma\, \delta_2(n)\, x^n = 4\left(\frac{x}{1-x} - \frac{x^3}{1-x^3} + \frac{x^5}{1-x^5} - \dots\right)$$
$$= 4\left(\frac{x}{1+x^2} + \frac{x^2}{1+x^4} + \frac{x^3}{1+x^6} + \dots\right).$$

It follows from § 10 that, if $s$ is a multiple of 4 then

$$(11{\cdot}11) \quad \delta_{2s}(n) = \frac{\pi^s n^{s-1}}{(s-1)!} \{1^{-s} c_1(n) + 2^{-s} c_4(n) + 3^{-s} c_3(n) + 4^{-s} c_8(n) + 5^{-s} c_5(n) + 6^{-s} c_{12}(n)$$
$$+ 7^{-s} c_7(n) + 8^{-s} c_{16}(n) + \dots\};$$

if $s$ is of the form $4k + 2$ then

$$(11{\cdot}21) \quad \delta_{2s}(n) = \frac{\pi^s n^{s-1}}{(s-1)!} \{1^{-s} c_1(n) - 2^{-s} c_4(n) + 3^{-s} c_3(n) - 4^{-s} c_8(n) + 5^{-s} c_5(n) - 6^{-s} c_{12}(n)$$
$$+ 7^{-s} c_7(n) - 8^{-s} c_{16}(n) + \dots\};$$

if $s$ is of the form $4k + 1$ then

$$(11{\cdot}31) \quad \delta_{2s}(n) = \frac{\pi^s n^{s-1}}{(s-1)!} \{1^{-s} c_1(n) + 2^{-s} s_4(n) - 3^{-s} c_3(n) + 4^{-s} s_8(n) + 5^{-s} c_5(n) + 6^{-s} s_{12}(n)$$
$$- 7^{-s} c_7(n) + 8^{-s} s_{16}(n) + \dots\},$$

except when $s = 1$; and if $s$ is of the form $4k + 3$ then

$$(11{\cdot}41) \quad \delta_{2s}(n) = \frac{\pi^s n^{s-1}}{(s-1)!} \{1^{-s} c_1(n) - 2^{-s} s_4(n) - 3^{-s} c_3(n) - 4^{-s} s_8(n) + 5^{-s} c_5(n) - 6^{-s} s_{12}(n)$$
$$- 7^{-s} c_7(n) - 8^{-s} s_{16}(n) + \dots\}.$$

* *Transactions of the Cambridge Philosophical Society*, vol. 22, 1916, pp. 159—184.

† For a more precise result concerning the order of $e_{2s}(n)$ see § 15.

From (11·5) and the remarks at the end of the previous section, it follows that

(11·51) $$r_2(n) = \delta_2(n) = \pi\{c_1(n) - \tfrac{1}{3}c_3(n) + \tfrac{1}{5}c_5(n) - \ldots\}$$
$$= \pi\{\tfrac{1}{2}s_4(n) + \tfrac{1}{4}s_8(n) + \tfrac{1}{6}s_{12}(n) + \ldots\},$$

but this is of course not such an elementary result as the preceding ones.

We can combine all the formulae (11·11)—(11·41) in one by writing

(11·6) $$\delta_{2s}(n) = \frac{\pi^s n^{s-1}}{(s-1)!}\{1^{-s}\mathbf{c}_1(n) + 2^{-s}\mathbf{c}_4(n) + 3^{-s}\mathbf{c}_3(n) + 4^{-s}\mathbf{c}_8(n) + 5^{-s}\mathbf{c}_5(n)$$
$$+ 6^{-s}\mathbf{c}_{12}(n) + 7^{-s}\mathbf{c}_7(n) + 8^{-s}\mathbf{c}_{16}(n) + \ldots,$$

where $s$ is an integer greater than 1 and

$$\mathbf{c}_r(n) = c_r(n)\cos\tfrac{1}{2}\pi s(r-1) - s_r(n)\sin\tfrac{1}{2}\pi s(r-1).$$

12. We can obtain analogous results concerning the number of representations of a number as the sum of 2, 4, 6, 8, ... triangular numbers. Equation (147) of my former paper* is equivalent to

(12·1) $$(1 - 2x + 2x^4 - 2x^9 + \ldots)^{2s} = 1 + \sum_1^\infty \delta_{2s}(n)(-x)^n + \frac{f^{4s}(x)}{f^{2s}(x^2)} \sum_{1 \leqslant n \leqslant \frac{1}{4}(s-1)} K_n(-x)^n \frac{f^{24n}(x^2)}{f^{24n}(x)},$$

where $K_n$ is a constant and

$$f(x) = (1-x)(1-x^2)(1-x^3)\ldots.$$

Suppose now that

$$x = e^{-\pi\alpha}, \quad x' = e^{-2\pi\alpha}.$$

Then we know that

(12·2) $$\sqrt{\alpha}\,(1 - 2x + 2x^4 - 2x^9 + \ldots) = 2x'^{\frac{1}{8}}(1 + x' + x'^3 + x'^6 + \ldots),$$

(12·3) $$\sqrt{(\tfrac{1}{2}\alpha)}\, x^{\frac{1}{24}} f(x) = x'^{\frac{1}{12}} f(x'^2), \quad \sqrt{\alpha}\, x^{\frac{1}{12}} f(x^2) = x'^{\frac{1}{24}} f(x').$$

Finally $1 + \sum_1^\infty \delta_{2s}(n)(-x)^n$ can be expressed in powers of $x'$ by using the formulae :—

(12·4) $$\alpha^s\left\{\tfrac{1}{2}\zeta(1-2s) + \frac{1^{2s-1}}{e^{2\alpha}-1} + \frac{2^{2s-1}}{e^{4\alpha}-1} - \frac{3^{2s-1}}{e^{6\alpha}-1} + \ldots\right\}$$
$$= (-\beta)^s\left\{\tfrac{1}{2}\zeta(1-2s) + \frac{1^{2s-1}}{e^{2\beta}-1} + \frac{2^{2s-1}}{e^{4\beta}-1} + \frac{3^{2s-1}}{e^{6\beta}-1} + \ldots\right\},$$

where $\alpha\beta = \pi^2$ and $s$ is an integer greater than 1 ; and

(12·5) $$(2\alpha)^{s+\frac{1}{2}}\left\{\frac{1^{2s}}{e^{\alpha}+e^{-\alpha}} + \frac{2^{2s}}{e^{2\alpha}+e^{-2\alpha}} + \frac{3^{2s}}{e^{3\alpha}+e^{-3\alpha}} + \ldots\right\}$$
$$= (-\beta)^s\sqrt{(2\beta)}\left\{\tfrac{1}{2}\eta(-2s) + \frac{1^{2s}}{e^{\beta}-1} - \frac{3^{2s}}{e^{3\beta}-1} + \frac{5^{2s}}{e^{5\beta}-1} - \ldots\right\},$$

where $\alpha\beta = \pi^2$, $s$ is any positive integer, and $\eta(s)$ is the function represented by the series $1^{-s} - 3^{-s} + 5^{-s} - \ldots$ and its analytical continuations.

It follows from all these formulae that, if $s$ is a positive integer and

(12·6) $$(1 + x + x^3 + x^6 + \ldots)^{2s} = \Sigma r'_{2s}(n)x^n = \Sigma\delta'_{2s}(n)x^n + \Sigma e'_{2s}(n)x^n,$$

then

$$\Sigma e'_{2s}(n)x^n = \frac{f^{4s}(x^2)}{f^{2s}(x)} \sum_{1 \leqslant n \leqslant \frac{1}{4}(s-1)} K_n(-x)^{-n}\frac{f^{24n}(x)}{f^{24n}(x^2)},$$

where $K_n$ and $f(x)$ are the same as in (12·1);

* *l.c.*, p. 181.

(12·61) $$(1^{-s}+3^{-s}+5^{-s}+\dots)\,\Sigma\,\delta'_{2s}(n)\,x^n=\frac{(\frac{1}{2}\pi)^s}{(s-1)!}\,x^{-\frac{1}{4}s}\left(\frac{1^{s-1}x}{1-x^2}+\frac{2^{s-1}x^2}{1-x^4}+\frac{3^{s-1}x^3}{1-x^6}+\dots\right)$$

if $s$ is a multiple of 4;

(12·62) $$(1^{-s}+3^{-s}+5^{-s}+\dots)\,\Sigma\,\delta'_{2s}(n)\,x^n=\frac{2(\frac{1}{4}\pi)^s}{(s-1)!}\,x^{-\frac{1}{4}s}\left(\frac{1^{s-1}x^{\frac{1}{2}}}{1-x}+\frac{3^{s-1}x^{\frac{3}{2}}}{1-x^3}+\frac{5^{s-1}x^{\frac{5}{2}}}{1-x^5}+\dots\right)$$

if $s$ is of the form $4k+2$;

(12·63) $$(1^{-s}-3^{-s}+5^{-s}-\dots)\,\Sigma\,\delta'_{2s}(n)\,x^n=\frac{2(\frac{1}{8}\pi)^s}{(s-1)!}\,x^{-\frac{1}{4}s}\left(\frac{1^{s-1}x^{\frac{1}{4}}}{1+x^{\frac{1}{2}}}+\frac{3^{s-1}x^{\frac{3}{4}}}{1+x^{\frac{3}{2}}}+\frac{5^{s-1}x^{\frac{5}{4}}}{1+x^{\frac{5}{2}}}+\dots\right.$$
$$\left.+\frac{1^{s-1}x^{\frac{1}{4}}}{1-x^{\frac{1}{2}}}-\frac{3^{s-1}x^{\frac{3}{4}}}{1-x^{\frac{3}{2}}}+\frac{5^{s-1}x^{\frac{5}{4}}}{1-x^{\frac{5}{2}}}-\dots\right)$$

if $s$ is of the form $4k+1$ (except when $s=1$); and

(12·64) $$(1^{-s}-3^{-s}+5^{-s}-\dots)\,\Sigma\,\delta'_{2s}(n)\,x^n=\frac{2(\frac{1}{8}\pi)^s}{(s-1)!}\,x^{-\frac{1}{4}s}\left(\frac{1^{s-1}x^{\frac{1}{4}}}{1+x^{\frac{1}{2}}}+\frac{3^{s-1}x^{\frac{3}{4}}}{1+x^{\frac{3}{2}}}+\frac{5^{s-1}x^{\frac{5}{4}}}{1+x^{\frac{5}{2}}}+\dots\right.$$
$$\left.-\frac{1^{s-1}x^{\frac{1}{4}}}{1-x^{\frac{1}{2}}}+\frac{3^{s-1}x^{\frac{3}{4}}}{1-x^{\frac{3}{2}}}-\frac{5^{s-1}x^{\frac{5}{4}}}{1-x^{\frac{5}{2}}}+\dots\right)$$

if $s$ is of the form $4k+3$. In the case in which $s=1$ we have

(12·65) $$\Sigma\,\delta_2'(n)\,x^n=x^{-\frac{1}{4}}\left(\frac{x^{\frac{1}{4}}}{1+x^{\frac{1}{2}}}+\frac{x^{\frac{3}{4}}}{1+x^{\frac{3}{2}}}+\frac{x^{\frac{5}{4}}}{1+x^{\frac{5}{2}}}+\dots\right)$$
$$=x^{-\frac{1}{4}}\left(\frac{x^{\frac{1}{4}}}{1-x^{\frac{1}{2}}}-\frac{x^{\frac{3}{4}}}{1-x^{\frac{3}{2}}}+\frac{x^{\frac{5}{4}}}{1-x^{\frac{5}{2}}}-\dots\right).$$

It is easy to see that the principal results proved about $e_{2s}(n)$ in my former paper are also true of $e'_{2s}(n)$, and in particular that

$$e'_{2s}(n)=0$$

when $s=1$, 2, 3 or 4, and

$$r'_{2s}(n)\sim\delta'_{2s}(n)$$

for all values of $s$.

13. It follows from (12·62) that if $s$ is of the form $4k+2$ then $(1^{-s}+3^{-s}+5^{-s}+\dots)\,\delta'_{2s}(n)$ is the coefficient of $x^n$ in

(13·1) $$\frac{2(\frac{1}{4}\pi)^s}{(s-1)!}\,x^{-\frac{1}{4}s}\left(\frac{1^{s-1}x^{\frac{1}{2}}}{1-x}+\frac{2^{s-1}x}{1-x^2}+\frac{3^{s-1}x^{\frac{3}{2}}}{1-x^3}+\dots\right).$$

Similarly from (12·63) and (12·64) it follows that if $s$ is an odd integer greater than 1 then $(1^{-s}-3^{-s}+5^{-s}-\dots)\,\delta'_{2s}(n)$ is the coefficient of $x^n$ in

(13·2) $$\frac{4(\frac{1}{8}\pi)^s}{(s-1)!}\,x^{-\frac{1}{4}s}\left(\frac{1^{s-1}x^{\frac{1}{4}}}{1+x^{\frac{1}{2}}}+\frac{2^{s-1}x^{\frac{1}{2}}}{1+x}+\frac{3^{s-1}x^{\frac{3}{4}}}{1+x^{\frac{3}{2}}}+\dots\right).$$

Now by applying our main formulae to (12·61), (13·1) and (13·2) we obtain:—

(13·3) $$\delta'_{2s}(n)=\frac{(\frac{1}{2}\pi)^s}{(s-1)!}\,(n+\tfrac{1}{4}s)^{s-1}\{1^{-s}c_1(n+\tfrac{1}{4}s)+3^{-s}c_3(n+\tfrac{1}{4}s)+5^{-s}c_5(n+\tfrac{1}{4}s)+\dots\}$$

if $s$ is a multiple of 4;

$$(13\cdot4)\qquad \delta'_{2s}(n)=\frac{(\frac{1}{2}\pi)^s}{(s-1)!}(n+\tfrac{1}{4}s)^{s-1}\{1^{-s}c_1(2n+\tfrac{1}{2}s)+3^{-s}c_3(2n+\tfrac{1}{2}s)+5^{-s}c_5(2n+\tfrac{1}{2}s)+\ldots$$

if $s$ is twice an odd number; and

$$(13\cdot5)\qquad \delta'_{2s}(n)=\frac{(\frac{1}{2}\pi)^s}{(s-1)!}(n+\tfrac{1}{4}s)^{s-1}\{1^{-s}c_1(4n+s)-3^{-s}c_3(4n+s)+5^{-s}c_5(4n+s)-\ldots\}$$

if $s$ is an odd number greater than 1.

Since the coefficient of $x^n$ in $(1+x+x^3+\ldots)^2$ is that of $x^{4n+1}$ in $(\frac{1}{2}+x+x^4+\ldots)^2$, it follows from (11·51) that

$$(13\cdot6)\qquad r_2'(n)=\delta_2'(n)=\frac{\pi}{4}\{c_1(4n+1)-\tfrac{1}{3}c_3(4n+1)+\tfrac{1}{5}c_5(4n+1)-\ldots\}.$$

This result however depends on the fact that the Dirichlet's series for $1/\eta(s)$ is convergent when $s=1$.

14. The preceding formulae for $\sigma_s(n)$, $\delta_{2s}(n)$, $\delta'_{2s}(n)$ may be arrived at by another method. We understand by

$$(14\cdot1)\qquad \frac{\sin n\pi}{k\sin(n\pi/k)}$$

the limit of

$$\frac{\sin x\pi}{k\sin(x\pi/k)}$$

when $x\to n$. It is easy to see that, if $n$ and $k$ are positive integers, and $k$ odd, then (14·1) is equal to 1 if $k$ is a divisor of $n$ and to 0 otherwise.

When $k$ is even we have (with similar conventions)

$$(14\cdot2)\qquad \frac{\sin n\pi}{k\tan(n\pi/k)}=1 \textit{ or } 0$$

according as $k$ is a divisor of $n$ or not. It follows that

$$(14\cdot3)\qquad \sigma_{s-1}(n)=n^{s-1}\left\{1^{-s}\left(\frac{\sin n\pi}{\sin n\pi}\right)+2^{-s}\left(\frac{\sin n\pi}{\tan\frac{1}{2}n\pi}\right)+3^{-s}\left(\frac{\sin n\pi}{\sin\frac{1}{3}n\pi}\right)+4^{-s}\left(\frac{\sin n\pi}{\tan\frac{1}{4}n\pi}\right)+\ldots\right\}.$$

Similarly from the definitions of $\delta_{2s}(n)$ and $\delta'_{2s}(n)$ we find that

$$(14\cdot4)\qquad \{1^{-s}+(-3)^{-s}+5^{-s}+(-7)^{-s}+\ldots\}\,\delta_{2s}(n)=\frac{\pi^s n^{s-1}}{(s-1)!}\left\{1^{-s}\left(\frac{\sin n\pi}{\sin n\pi}\right)\right.$$
$$\left.+2^{-s}\left(\frac{\sin n\pi}{\sin(\frac{1}{2}n\pi+\frac{1}{2}s\pi)}\right)+3^{-s}\left(\frac{\sin n\pi}{\sin(\frac{1}{3}n\pi+s\pi)}\right)+4^{-s}\left(\frac{\sin n\pi}{\sin(\frac{1}{4}n\pi+\frac{3}{2}s\pi)}\right)+\ldots\right\}$$

if $s$ is an integer greater than 1;

$$(14\cdot5)\qquad r_2(n)=\delta_2(n)=4\left\{\left(\frac{\sin n\pi}{\sin n\pi}\right)-\tfrac{1}{3}\left(\frac{\sin n\pi}{\sin\frac{1}{3}n\pi}\right)+\tfrac{1}{5}\left(\frac{\sin n\pi}{\sin\frac{1}{5}n\pi}\right)-\ldots\right\}$$
$$=4\left\{\tfrac{1}{2}\left(\frac{\sin n\pi}{\cos\frac{1}{2}n\pi}\right)-\tfrac{1}{4}\left(\frac{\sin n\pi}{\cos\frac{1}{4}n\pi}\right)+\tfrac{1}{6}\left(\frac{\sin n\pi}{\cos\frac{1}{6}n\pi}\right)-\ldots\right\};$$

$$(14\cdot6)\qquad (1^{-s}+3^{-s}+5^{-s}+\ldots)\,\delta'_{2s}(n)=\frac{(\frac{1}{2}\pi)^s}{(s-1)!}(n+\tfrac{1}{4}s)^{s-1}\left\{1^{-s}\left(\frac{\sin(n+\frac{1}{4}s)\pi}{\sin(n+\frac{1}{4}s)\pi}\right)\right.$$
$$\left.+3^{-s}\left(\frac{\sin(n+\frac{1}{4}s)\pi}{\sin\frac{1}{3}(n+\frac{1}{4}s)\pi}\right)+5^{-s}\left(\frac{\sin(n+\frac{1}{4}s)\pi}{\sin\frac{1}{5}(n+\frac{1}{4}s)\pi}\right)+\ldots\right\}$$

if $s$ is a multiple of 4;

$$(14\cdot7)\quad (1^{-s}+3^{-s}+5^{-s}+\dots)\,\delta'_{2s}(n)=\frac{(\frac{1}{2}\pi)^s}{(s-1)!}(n+\tfrac{1}{4}s)^{s-1}\left\{1^{-s}\left(\frac{\sin(2n+\frac{1}{2}s)\pi}{\sin(2n+\frac{1}{2}s)\pi}\right)+3^{-s}\left(\frac{\sin(2n+\frac{1}{2}s)\pi}{\sin\frac{1}{3}(2n+\frac{1}{2}s)\pi}\right)+5^{-s}\left(\frac{\sin(2n-\frac{1}{2}s)\pi}{\sin\frac{1}{5}(2n+\frac{1}{2}s)\pi}\right)+\dots\right\}$$

if $s$ is twice an odd number;

$$(14\cdot8)\quad (1^{-s}-3^{-s}+5^{-s}-\dots)\,\delta'_{2s}(n)=\frac{(\frac{1}{2}\pi)^s}{(s-1)!}(n+\tfrac{1}{4}s)^{s-1}\left\{1^{-s}\left(\frac{\sin(4n+s)\pi}{\sin(4n+s)\pi}\right)-3^{-s}\left(\frac{\sin(4n+s)\pi}{\sin\frac{1}{3}(4n+s)\pi}\right)+5^{-s}\left(\frac{\sin(4n+s)\pi}{\sin\frac{1}{5}(4n+s)\pi}\right)-\dots\right\}$$

if $s$ is an odd number greater than 1; and

$$(14\cdot9)\quad r'(n)=\delta_2(n)=\left\{\left(\frac{\sin(4n+1)\pi}{\sin(4n+1)\pi}\right)-\tfrac{1}{3}\left(\frac{\sin(4n+1)\pi}{\sin\frac{1}{3}(4n+1)\pi}\right)+\tfrac{1}{5}\left(\frac{\sin(4n+1)\pi}{\sin\frac{1}{5}(4n+1)\pi}\right)-\dots\right\}.$$

In all these equations the series on the right hand are finite Dirichlet's series and therefore absolutely convergent.

But the series (14·3) is (as is easily shown by actual multiplication) the product of the two series

$$1^{-s}c_1(n)+2^{-s}c_2(n)+\dots$$

and

$$n^{s-1}(1^{-s}+2^{-s}+3^{-s}+\dots).$$

We thus obtain an alternative proof of the formulae (7·5). Similarly taking the previous expression of $\delta_{2s}(n)$, viz. the right-hand side of (11·6), and multiplying it by the series

$$1^{-s}+(-3)^{-s}+5^{-s}+(-7)^{-s}+\dots$$

we can show that the product is actually the right-hand side of (14·4). The formulae for $\delta'_{2s}(n)$ can be disposed of similarly.

15. The formulae which I have found are closely connected with a method used for another purpose by Mr Hardy and myself*. The function

$$(15\cdot1)\qquad (1+2x+2x^4+2x^9+\dots)^s=\Sigma\, r_{2s}(n)\,x^n$$

has every point of the unit circle as a singular point. If $x$ approaches a 'rational point' $\exp(-2p\pi i/q)$ on the circle, the function behaves roughly like

$$(15\cdot2)\qquad \frac{\pi^s(\omega_{p,q})^s}{\{-(2p\pi i/q)-\log x\}^s}$$

where $\omega_{p,q}=1$, 0, or $-1$ according as $q$ is of the form $4k+1$, $4k+2$ or $4k+3$, while if $q$ is of the form $4k$ then $\omega_{p,q}=-2i$ or $2i$ according as $p$ is of the form $4k+1$ or $4k+3$.

Following the argument of our paper referred to, we can construct simple functions of $x$ which are regular except at one point of the circle of convergence, and there behave in a manner very similar to that of the function (15·1): for example at the point $\exp(-2p\pi i/q)$ such a function is

$$(15\cdot3)\qquad \frac{\pi^s(\omega_{p,q})^s}{(s-1)!}\sum_1^\infty n^{s-1}e^{2np\pi i/q}x^n.$$

* 'Asymptotic formulae in Combinatory Analysis', *Proc. London Math. Soc.*, ser. 2, vol. 17, 1918, pp. 75—115.

The method which we used, with particular reference to the function

(15·4) $$\frac{1}{(1-x)(1-x^2)(1-x^3)\dots} = \Sigma\, p(n)\, x^n,$$

was to approximate to the coefficients by means of a sum of a large number of the coefficients of these auxiliary functions. This method leads, in the present problem, to formulae of the type

$$r_{2s}(n) = \delta_{2s}(n) + O(n^{\frac{1}{2}s}),$$

the first term on the right-hand side presenting itself precisely in the form of the series (11·11) etc.

It is a very interesting problem to determine in such cases whether the approximate formula gives an exact representation of such an arithmetical function. The results proved here show that, in the case of $r_{2s}(n)$ this is in general not so. The formula represents not $r_{2s}(n)$ but (except when $s=1$) its dominant term $\delta_{2s}(n)$, which is equal to $r_{2s}(n)$ only when $s=1, 2, 3$, or $4$. When $s=1$ the formula gives $2\delta_2(n)$*.

16. We shall now consider the sum

(16·1) $$\sigma_s(1) + \sigma_s(2) + \dots + \sigma_s(n).$$

Suppose that

(16·2) $$T_r(n) = \tfrac{1}{2}\sum_{\lambda}\left(\frac{\sin\{(2n+1)\pi\lambda/r\}}{\sin(\pi\lambda/r)} - 1\right), \qquad U_r(n) = \tfrac{1}{2}\sum_{\lambda}\frac{\sin\{(2n+1)\pi\lambda/r\}}{\sin(\pi\lambda/r)},$$

where $\lambda$ is prime to $r$ and does not exceed $r$, so that

$$T_r(n) = c_r(1) + c_r(2) + \dots + c_r(n)$$

and

$$U_r(n) = T_r(n) + \tfrac{1}{2}\phi(r),$$

where $\phi(n)$ is the same as in § 9. Since $c_r(n) = O(1)$ as $r \to \infty$, it follows that

(16·21) $$T_r(n) = O(1), \qquad U_r(n) = O(r),$$

as $r\to\infty$. It follows from (7·5) that if $s>0$ then

(16·3) $$\sigma_{-s}(1) + \sigma_{-s}(2) + \dots + \sigma_{-s}(n) = \zeta(s+1)\left\{n + \frac{T_2(n)}{2^{s+1}} + \frac{T_3(n)}{3^{s+1}} + \frac{T_4(n)}{4^{s+1}} + \dots\right\}.$$

Since

$$\sum_1^\infty \frac{\phi(\nu)}{\nu^{s+1}} = \frac{\zeta(s)}{\zeta(s+1)}$$

if $s>1$, (16·3) can be written as

(16·31) $$\sigma_{-s}(1) + \sigma_{-s}(2) + \dots + \sigma_{-s}(n)$$
$$= \zeta(s+1)\left\{n + \tfrac{1}{2} + \frac{U_2(n)}{2^{s+1}} + \frac{U_3(n)}{3^{s+1}} + \frac{U_4(n)}{4^{s+1}} + \dots\right\} - \tfrac{1}{2}\zeta(s),$$

if $s>1$. Similarly from (8·3), (8·4) and (11·51) we obtain

(16·4) $$d(1) + d(2) + \dots + d(n) = -\tfrac{1}{2}T_2(n)\log 2 - \tfrac{1}{3}T_3(n)\log 3 - \tfrac{1}{4}T_4(n)\log 4 - \dots,$$

(16·5) $$d(1)\log 1 + d(2)\log 2 + \dots + d(n)\log n$$
$$= \tfrac{1}{2}T_2(n)\{2\nu\log 2 - (\log 2)^2\} + \tfrac{1}{3}T_3(n)\{2\nu\log 3 - (\log 3)^2\} + \dots,$$

(16·6) $$r_2(1) + r_2(2) + \dots + r_2(n) = \pi\{n - \tfrac{1}{3}T_3(n) + \tfrac{1}{5}T_5(n) - \tfrac{1}{7}T_7(n) + \dots\}.$$

* The method is also applicable to the problem of the representation of a number by the sum of an *odd* number of squares, and gives an exact result when the number of squares is 3, 5, or 7. See G. H. Hardy, 'On the representation of a number as the sum of any number of squares, and in particular of five or seven', *Proc. London Math. Soc.* (*Records of proceedings at meetings*, March 1918). A fuller account of this paper will appear shortly in the *Proceedings of the National Academy of Sciences* (Washington, D.C.).

Suppose now that

$$T_{r,s}(n) = \sum_{\lambda}\left(1^s \cos\frac{2\pi\lambda}{r} + 2^s \cos\frac{4\pi\lambda}{r} + \ldots + n^s \cos\frac{2n\pi\lambda}{r}\right),$$

where $\lambda$ is prime to $r$ and does not exceed $r$, so that

$$T_{r,s}(n) = 1^s c_r(1) + 2^s c_r(2) + \ldots + n^s c_r(n).$$

Then it follows from (7·5) that

$$(16{\cdot}7)\quad \sigma_s(1) + \sigma_s(2) + \ldots + \sigma_s(n)$$
$$= \zeta(s+1)\left\{(1^s + 2^s + \ldots + n^s) + \frac{T_{2,s}(n)}{2^{s+1}} + \frac{T_{3,s}(n)}{3^{s+1}} + \frac{T_{4,s}(n)}{4^{s+1}} + \ldots\right\}$$

if $s > 0$. Putting $s = 1$ in (16·3) and (16·7), we find that

$$(16{\cdot}8)\quad (n-1)\sigma_{-1}(1) + (n-2)\sigma_{-1}(2) + \ldots + (n-n)\sigma_{-1}(n)$$
$$= \frac{\pi^2}{6}\left\{\frac{n(n-1)}{2} + \frac{\nu_2(n)}{2^2} + \frac{\nu_3(n)}{3^2} + \frac{\nu_4(n)}{4^2} + \ldots\right\},$$

where

$$\nu_r(n) = \tfrac{1}{2}\sum_{\lambda}\left\{\frac{\sin^2(\pi n\lambda/r)}{\sin^2(\pi\lambda/r)} - n\right\},$$

$\lambda$ being prime to $r$ and not exceeding $r$.

It has been proved by Wigert*, by less elementary methods, that the left-hand side of (16·8) is equal to

$$(16{\cdot}9)\quad \frac{\pi^2}{12}n^2 - \tfrac{1}{2}n(\gamma - 1 + \log 2n\pi) - \tfrac{1}{24} + \frac{\sqrt{n}}{2\pi}\sum_{1}^{\infty}\frac{\sigma_{-1}(\nu)}{\sqrt{\nu}} J_1\{4\pi\sqrt{(\nu n)}\},$$

where $J_1$ is the ordinary Bessel's function.

17. We shall now find a relation between the functions (16·1) and (16·3) which enables us to determine the behaviour of the former for large values of $n$. It is easily shown that this function is equal to

$$(17{\cdot}1)\quad \sum_{\nu=1}^{\sqrt{n}}\left(1^s + 2^s + 3^s + \ldots + \left[\frac{n}{\nu}\right]^s\right) + \sum_{\nu=1}^{\sqrt{n}}\nu^s\left[\frac{n}{\nu}\right] - [\sqrt{n}]\sum_{\nu=1}^{\sqrt{n}}\nu^s.$$

Now

$$1^s + 2^s + \ldots + k^s = \zeta(-s) + \frac{(k+\frac{1}{2})^{s+1}}{s+1} + O(k^{s-1})$$

for all values of $s$, it being understood that

$$\zeta(-s) + \frac{(k+\frac{1}{2})^{s+1}}{s+1}$$

denotes $\gamma + \log(k+\frac{1}{2})$ when $s = -1$. Let

$$\left[\frac{n}{\nu}\right] = \frac{n}{\nu} - \tfrac{1}{2} + \epsilon_\nu, \qquad [\sqrt{n}] = t = \sqrt{n} - \tfrac{1}{2} + \epsilon.$$

Then we have

$$1^s + 2^s + \ldots + \left[\frac{n}{\nu}\right]^s = \zeta(-s) + \frac{1}{s+1}\left(\frac{n}{\nu}\right)^{s+1} + \epsilon_\nu\left(\frac{n}{\nu}\right)^s + O\left(\frac{n^{s-1}}{\nu^{s-1}}\right)$$

and

$$\nu^s\left[\frac{n}{\nu}\right] = n\nu^{s-1} - \tfrac{1}{2}\nu^s + \epsilon_\nu \nu^s.$$

* *Acta Mathematica*, vol. 37, pp. 113–140 (p. 140).

It follows from these equations and (17·1) that

$$(17\cdot2)\qquad \sigma_s(1)+\sigma_s(2)+\ldots+\sigma_s(n)=\sum_{\nu=1}^{t}\Big\{\zeta(-s)+\frac{1}{s+1}\left(\frac{n}{\nu}\right)^{s+1}+n\nu^{s-1}$$
$$+\epsilon_\nu\left(\frac{n}{\nu}\right)^{s}+\epsilon_\nu\nu^{s}-(\sqrt{n}+\epsilon)\nu^{s}+O\left(\frac{n^{s-1}}{\nu^{s-1}}\right)\Big\}.$$

Changing $s$ to $-s$ in (17·2) we have

$$(17\cdot21)\qquad n^s\{\sigma_{-s}(1)+\sigma_{-s}(2)+\ldots+\sigma_{-s}(n)\}=\sum_{\nu=1}^{t}\Big\{n^s\zeta(s)+\frac{n\nu^{s-1}}{1-s}+\left(\frac{n}{\nu}\right)^{s+1}+\epsilon_\nu\nu^{s}$$
$$+\epsilon_\nu\left(\frac{n}{\nu}\right)^{s}-(\sqrt{n}+\epsilon)\left(\frac{n}{\nu}\right)^{s}+O\left(\frac{\nu^{s+1}}{n}\right)\Big\}.$$

It follows that

$$(17\cdot3)\qquad n^s\{\sigma_{-s}(1)+\sigma_{-s}(2)+\ldots+\sigma_{-s}(n)\}-\{\sigma_s(1)+\sigma_s(2)+\ldots+\sigma_s(n)\}=\sum_{\nu=1}^{t}\Big\{n^s\zeta(s)-\zeta(-s)$$
$$+\frac{s}{1+s}\left(\frac{n}{\nu}\right)^{s+1}+\frac{s}{1-s}n\nu^{s-1}+(\sqrt{n}+\epsilon)\nu^{s}-(\sqrt{n}+\epsilon)\left(\frac{n}{\nu}\right)^{s}+O\left(\frac{n^{s-1}}{\nu^{s-1}}+\frac{\nu^{s+1}}{n}\right)\Big\}.$$

Suppose now that $s>0$. Then, since $\nu$ varies from 1 to $t$, it is obvious that

$$\frac{\nu^{s+1}}{n}<\frac{n^{s-1}}{\nu^{s-1}}$$

and so

$$O\left(\frac{\nu^{s+1}}{n}\right)=O\left(\frac{n^{s-1}}{\nu^{s-1}}\right).$$

Also

$$\sum_{\nu=1}^{t}\{n^s\zeta(s)-\zeta(-s)\}=(\sqrt{n}-\tfrac{1}{2}+\epsilon)\{n^s\zeta(s)-\zeta(-s)\};$$

$$\sum_{\nu=1}^{t}\frac{s}{1+s}\left(\frac{n}{\nu}\right)^{s+1}=\frac{sn^{s+1}}{1+s}\zeta(1+s)-\frac{n^{s+1}}{s+1}(\sqrt{n}+\epsilon)^{-s}+O(n^{\frac{1}{2}s});$$

$$\sum_{\nu=1}^{t}\frac{s}{1-s}n\nu^{s-1}=\frac{ns}{1-s}\zeta(1-s)+\frac{n}{1-s}(\sqrt{n}+\epsilon)^{s}+O(n^{\frac{1}{2}s});$$

$$\sum_{\nu=1}^{t}(\sqrt{n}+\epsilon)\nu^{s}=(\sqrt{n}+\epsilon)\zeta(-s)+\frac{(\sqrt{n}+\epsilon)^{2+s}}{1+s}+O(n^{\frac{1}{2}s});$$

$$\sum_{\nu=1}^{t}(\sqrt{n}+\epsilon)\left(\frac{n}{\nu}\right)^{s}=n^s(\sqrt{n}+\epsilon)\zeta(s)+\frac{n^s}{1-s}(\sqrt{n}+\epsilon)^{2-s}+O(n^{\frac{1}{2}s});$$

and

$$\sum_{\nu=1}^{t}O\left(\frac{n^{s-1}}{\nu^{s-1}}\right)=O(m),$$

where

$$(17\cdot4)\qquad m=n^{\frac{1}{2}s}\ (s<2),\quad m=n\log n\ (s=2),\quad m=n^{s-1}\ (s>2).$$

It follows that the right-hand side of (17·3) is equal to

$$\frac{sn^{1+s}}{1+s}\zeta(1+s)+\frac{sn}{1-s}\zeta(1-s)-\tfrac{1}{2}n^s\zeta(s)+\frac{(\sqrt{n}+\epsilon)^{2+s}-n^{s+1}(\sqrt{n}+\epsilon)^{-s}}{1+s}$$
$$+\frac{n(\sqrt{n}+\epsilon)^{s}-n^s(\sqrt{n}+\epsilon)^{2-s}}{1-s}+O(m).$$

But

$$\frac{(\sqrt{n}+\epsilon)^{2+s}-n^{s+1}(\sqrt{n}+\epsilon)^{-s}}{1+s}=2\epsilon n^{\frac{1}{2}(1+s)}+O(n^{\frac{1}{2}s}):$$

$$\frac{n(\sqrt{n}+\epsilon)^{s}-n^s(\sqrt{n}+\epsilon)^{2-s}}{1-s}=-2\epsilon n^{\frac{1}{2}(1+s)}+O(n^{\frac{1}{2}s}).$$

It follows that

$$(17\cdot5)\qquad \sigma_s(1)+\sigma_s(2)+\ldots+\sigma_s(n)=n^s\{\sigma_{-s}(1)+\sigma_{-s}(2)+\ldots+\sigma_{-s}(n)\}-\frac{sn^{1+s}}{1+s}\zeta(1+s)$$
$$+\tfrac{1}{2}n^s\zeta(s)-\frac{sn}{1-s}\zeta(1-s)+O(m)$$

if $s>0$, $m$ being the same as in (17·4). If $s=1$, (17·5) reduces to

$$(17\cdot6)\qquad (n-1)\sigma_{-1}(1)+(n-2)\sigma_{-1}(2)+\ldots+(n-n)\sigma_{-1}(n)$$
$$=\frac{\pi^2}{12}n^2-\tfrac{1}{2}n(\gamma-1+\log 2n\pi)+O(\sqrt{n})^*.$$

From (16·2) and (17·5) it follows that

$$(17\cdot7)\qquad \sigma_s(1)+\sigma_s(2)+\ldots+\sigma_s(n)=\frac{n^{1+s}}{1+s}\zeta(1+s)+\tfrac{1}{2}n^s\zeta(s)+\frac{sn}{s-1}\zeta(1-s)$$
$$+n^s\zeta(1+s)\left\{\frac{T_2(n)}{2^{s+1}}+\frac{T_3(n)}{3^{s+1}}+\frac{T_4(n)}{4^{s+1}}+\ldots\right\}+O(m),$$

for all positive values of $s$. If $s>1$, the right-hand side can be written as

$$(17\cdot8)\qquad \frac{ns}{s-1}\zeta(1-s)+n^s\zeta(1+s)\left\{\frac{n}{1+s}+\tfrac{1}{2}+\frac{U_2(n)}{2^{s-1}}+\frac{U_3(n)}{3^{s+1}}+\frac{U_4(n)}{4^{s+1}}+\ldots\right\}+O(m).$$

Putting $s=1$ in (17·7) we obtain

$$(17\cdot9)\qquad \sigma_1(1)+\sigma_1(2)+\ldots+\sigma_1(n)=\frac{\pi^2}{12}n^2+\tfrac{1}{2}n(\gamma-1+\log 2n\pi)$$
$$+\frac{\pi^2 n}{6}\left\{\frac{T_2(n)}{2^2}+\frac{T_3(n)}{3^2}+\frac{T_4(n)}{4^2}+\ldots\right\}+O(\sqrt{n}).$$

*Additional note to § 7 (May 1, 1918).*

From (7·2) it follows that

$$\frac{1}{\zeta(r)}\{1^{-s}\sigma_{1-r}(1)+2^{-s}\sigma_{1-r}(2)+\ldots\}=1^{-s}\sum_1^\infty m^{-r}c_m(1)+2^{-s}\sum_1^\infty m^{-r}c_m(2)+\ldots,$$

or

$$\frac{\zeta(s)\,\zeta(r+s-1)}{\zeta(r)}=\sum_1^\infty\sum_1^\infty\frac{c_m(n)}{m^r n^s},$$

from which we deduce

$$\zeta(s)\sum_\delta\mu(\delta)\,\delta'^{1-s}=\frac{c_m(1)}{1^s}+\frac{c_m(2)}{2^s}+\frac{c_m(3)}{3^s}+\ldots,$$

$\delta$ being a divisor of $m$ and $\delta'$ its conjugate. The series on the right-hand side is convergent for $s>0$ (except when $m=1$, when it reduces to the ordinary series for $\zeta(s)$).

When $s=1$, $m>1$, we have to replace the left-hand side by its limit as $s\to1$. We find that

$$c_m(1)+\tfrac{1}{2}c_m(2)+\tfrac{1}{3}c_m(3)+\ldots=-\Lambda(m),$$

$\Lambda(m)$ being the well-known arithmetical function which is equal to $\log p$ if $m$ is a power of a prime $p$ and to zero otherwise.

* This result has been proved by Landau. See his report on Wigert's memoir in the *Göttingische gelehrte Anzeigen*, 1915, pp. 377–414 (p. 402). Landau has also, by a more transcendental method, replaced $O(\sqrt{n})$ by $O(n^{\frac{2}{5}})$ (*l.c.* p. 414).

# XIV. *Asymptotic expansions of hypergeometric functions.*

By G. N. Watson, Sc.D., Trinity College, Cambridge.

[*Received* June 11, 1917. *Read* Feb. 4, 1918.]

1. The hypergeometric function $F(\alpha, \beta; \gamma; x)$ presents two distinct problems to mathematicians interested in the theories of analytic continuations and asymptotic expansions. The first, and simpler, problem is that of finding the analytic continuation of the function beyond the circle $|x| = 1$, which is the circle of convergence of the series by which the function is usually defined. More generally, the problem is that of finding the analytic continuation of ${}_{q+1}F_q$ beyond the circle $|x| = 1$, and of finding the asymptotic expansion of ${}_pF_q$ for large values of $|x|$ when $p < q+1$; as usual, ${}_pF_q$ denotes a generalised hypergeometric function defined by a series in which each coefficient is a fraction whose numerator and denominator consist of $p$ and of $q+1$ sets of factors respectively; the function is an integral function when $p < q+1$.

This problem has now been completely solved; the earlier investigations by double-circuit integrals (in connexion with which reference may be made to the researches of Hankel* on Bessel functions, of Hobson† on Legendre functions, and the extensions due to Orr‡, by means of elaborate inductions, to generalised hypergeometric functions) have been followed by the memoirs published by Barnes§, whose powerful method of employing integrals involving gamma functions renders it unlikely that the subject retains any general results to be discovered by future investigators.

The second problem presented by the hypergeometric function is that of the discovery of approximate formulae (and complete asymptotic expansions) for the function when one, or more, of the constants $\alpha, \beta, \gamma$ is large and the remaining constants and $x$ have any assigned values. The earliest investigation of a problem of this type seems to be due to Laplace‖, who gave two proofs that, when $n$ is a large integer and $0 < \theta < \pi$, then

$$P_n(\cos\theta) \sim \left(\frac{2}{n\pi \sin\theta}\right)^{\frac{1}{2}} \cos\{(n+\tfrac{1}{2})\theta - \tfrac{1}{4}\pi\}.$$

A more satisfactory demonstration of Laplace's result is given by Darboux¶ in his epoch-making memoirs *Sur l'approximation des fonctions de très grands nombres*. These memoirs also contain an investigation of the hypergeometric function $F(\alpha + n, -n; \gamma; x)$, where $n$ is a large positive integer; this function is sometimes known as Jacobi's** (or Tchebychef's††) polynomial.

* *Math. Ann.* I. (1869), pp. 467–501.

† *Phil. Trans.* 187 A (1896), pp. 443–531.

‡ *Cambridge Phil. Trans.* XVII. (1898), pp. 171–199, 283–290.

§ *Proc. London Math. Soc.* (2) V. (1907), pp. 59–118; VI. (1908), pp. 141–177: *Quarterly Journal* XXXIX. (1908), pp. 97–204.

‖ *Mécanique Céleste* V. (1823), livre 11, supplément 1.

¶ *Liouville* (3) IV. (1878), pp. 5–56, 377–416.

** *Crelle* LVI. (1859), pp. 149–175.

†† *Œuvres* II. (1907), pp. 189–215; these researches were first published in 1872–1874.

More modern investigations in the particular case of Legendre functions are due to Barnes*, but his methods convey the impression that they are primarily adapted for attacking the first problem rather than the second; on the other hand, it must be stated that they are quite effective in obtaining complete asymptotic expansions of $P_n^m(z)$ and $Q_n^m(z)$ when either $n$ or $m$ is large, and $n$, $m$ are not restricted to be integers.

The most natural way of attacking the second problem seems to me to be by the method of steepest descents. The application of this method to the problem is of special historical interest, because it was in connexion with hypergeometric functions that Riemann wrote the paper† which contains the first indications of the potentialities of the method. More recently, in the hands of Debye, Brillouin and myself, the method has proved to be effective in dealing with Bessel functions, Weber-Hermite functions (*i.e.* those associated in harmonic analysis with the parabolic cylinder) and numerous other functions (defined by definite integrals of various types) which occur in many branches of Mathematical Physics‡.

As the investigations of this paper have the asymptotic expansions of Legendre functions as one of their ultimate objects, a notation will be employed which will make it as easy as possible to write down the special results for these functions. It may be mentioned here that the contours which are yielded by the method of steepest descents in the case of hypergeometric functions are all algebraic curves, many of them being nodal circular cubics; this is in marked contradistinction to the fact that the contours employed in previous applications of the method (with the exception of some of Brillouin's researches on the functions of Physical Optics) have been somewhat complicated transcendental curves.

It should also be remarked that very slight modifications of the contour integrals are adequate to supply the asymptotic expansions in various exceptional cases in which the application of the Mellin-Barnes method requires detailed separate investigations.

* *Quarterly Journal, loc. cit.* This paper contains (p. 143) an extended account of researches on Legendre functions. For more general hypergeometric functions, see also *Encyclopédie des Sciences Math.* t. II. vol. 5.

Legendre functions have been discussed from the aspect of the theory of differential equations by Nicholson, *Quarterly Journal* XLI. (1910), pp. 241–264, XLIII. (1911), pp. 53–62.

It may also be noted here that a generalisation of Laplace's formula

$$P_n^m(\cos\theta) = \frac{2}{\sqrt{\pi}} \frac{\Gamma(n+m+1)}{\Gamma(n+\frac{3}{2})} \left[ \frac{\cos\{(n+\frac{1}{2})\theta - \frac{1}{4}\pi + \frac{1}{2}m\pi\}}{(2\sin\theta)^{\frac{1}{2}}} + \frac{1^2-4m^2}{2(2n+3)} \frac{\cos\{(n+\frac{3}{2})\theta - \frac{3}{4}\pi + \frac{1}{2}m\pi\}}{(2\sin\theta)^{\frac{3}{2}}} + \dots \right]$$

(Hobson, p. 486) affords a *convergent* expansion when $\frac{1}{6}\pi < \theta < \frac{5}{6}\pi$, which is *asymptotic* over the wider range $0 < \theta < \pi$, when $n$ is large. It will appear in Part IV of this memoir that a similar property characterises hypergeometric functions which are there described as belonging to type B.

† This paper, which was published posthumously, is given in Riemann's *Werke* (1892), pp. 424–430. See also the French edition (1898), pp. 369–377.

‡ Debye, *Math. Ann.* LXVII. (1909), pp. 535–568; *Münchener Sitzungsberichte* (1910), [5]. Brillouin, *Ann. de l'École Normale Supérieure* (3) XXXIII. (1916), pp. 17–69. Watson, *Proc. London Math. Soc.* (2) XVI. (1917), pp. 150–174, XVII. (1918), pp. 116–148; *Quarterly Journal* XLVIII. (1917), pp. 1–18.

## PART I. THE GENERALISED JACOBI-TCHEBYCHEF FUNCTIONS.

### 2. *Statement of the problem.*

In this section, we shall obtain the complete asymptotic expansion of

$$F(\alpha+\lambda, \beta-\lambda; \gamma; x),$$

where $\alpha, \beta, \gamma, x$ have any assigned values, real or complex (the value of $|x|$ being not restricted to be less than 1), and $|\lambda|$ is large while* $|\arg\lambda| < \pi$. More precisely it is supposed that $|\arg\lambda| \leqslant \pi - \delta$, where $\delta$ is an arbitrary positive number, independent of $|\lambda|$.

We shall invariably write $\frac{1}{2} - \frac{1}{2}z$ in place of $x$ (with a view to applications to Legendre functions†) and $\zeta$ will then be defined by the equation

$$z = \cosh\zeta;$$

and, if $\zeta = \xi + i\eta$, where $\xi, \eta$ are real, it will be supposed that $\xi \geqslant 0, -\pi \leqslant \eta \leqslant \pi$. These conventions determine $\zeta$ uniquely when $z$ is given, except when $z$ is real and less than 1. It will be supposed that the arguments of $z, z-1, z+1$ are given their principal values (numerically not exceeding $\pi$), and, in the special case in which $z-1$ is real and negative, it is to be supposed that $z$ attains its value by a limiting process which will determine whether $\arg(z-1)$ is to be taken equal to $+\pi$ or to $-\pi$. The values of $\arg z$ and $\arg(z+1)$ are to be determined in this exceptional case in the same manner.

In the $z$-plane the curves on which $\xi$ is constant are confocal ellipses, and the curves on which $\eta$ is constant are arcs of confocal hyperbolas, the foci being at $z = \pm 1$.

### 3. *The contour integrals associated with the hypergeometric function.*

With the notation introduced in § 2, the function under consideration is

$$F(\alpha+\lambda, \beta-\lambda; \gamma; -\sinh^2\tfrac{1}{2}\zeta) \equiv F(\alpha+\lambda, \beta-\lambda; \gamma; \tfrac{1}{2}-\tfrac{1}{2}z).$$

In our preliminary discussion we suppose that $|\arg\lambda| \leqslant \frac{1}{2}\pi - \delta$, and not $\pi - \delta$ as stated in § 2; the more extended range is considered in § 7.

To deal with this hypergeometric function we write down one of the integrals satisfying the associated hypergeometric equation, namely‡

$$I \equiv \int_{-1}^{1} (1-t)^{[illegible]} (1+t)^{[illegible]} (z-t)^{[illegible]} dt,$$

in which the path of integration is either (i) the segment of the real axis, or (ii) a path which is reconcilable with this segment without crossing over the singularity $t = z$; and

* Some of the expansions which will be obtained are valid only when $-\frac{1}{2}\pi - \omega_2 < \arg\lambda < \frac{1}{2}\pi - \omega_1$, where $\omega_1, \omega_2$ are certain positive acute angles. The asymptotic expansions when $\lambda$ is not in this sector of the plane are obtained by working with $\lambda_1 (\equiv -\lambda)$ and making an obvious interchange of $\alpha, \beta$.

† Of course $\xi$ has no connexion with the real part of $z$. The reader is doubtless familiar with formulae of the type

$$P_n^m(z) = \frac{1}{\Gamma(1-m)}\left(\frac{z+1}{z-1}\right)^{\frac{1}{2}m} F(n+1, -n; 1-m; \tfrac{1}{2}-\tfrac{1}{2}z).$$

‡ This is easily derived from the integral given by Forsyth, *Differential Equations*, § 143.

$\arg(z-t)$ is defined by giving it the value $\arg(z+1)$ at $t=-1$. The integral obviously converges at both limits when $\lambda$ is sufficiently large.

To evaluate the integral, suppose that $|z-1|>2$; and then, taking the path of integration to be the real axis, we have $|1-t|<|z-1|$. Hence, on expansion, we get

$$I_1=(z-1)^{-\alpha-\lambda}\int_{-1}^{1}(1-t)^{\alpha+\lambda-\gamma}(1+t)^{\gamma-\beta+\lambda-1}\sum_{n=0}^{\infty}\frac{\Gamma(\alpha+\lambda+n)}{n!\,\Gamma(\alpha+\lambda)}\left(\frac{1-t}{1-z}\right)^{n}dt.$$

Integrating term by term, we have at once

$$I_1=\frac{2^{\lambda-\beta}\,\Gamma(\alpha+\lambda-\gamma+1)\,\Gamma(\gamma-\beta+\lambda)}{\Gamma(\alpha-\beta+2\lambda+1)}\left(\frac{z-1}{2}\right)^{-\alpha-\lambda}F\left(\alpha+\lambda,\ \alpha+\lambda-\gamma+1;\ \alpha-\beta+2\lambda+1;\ \frac{2}{1-z}\right).$$

Since $I_1$ is analytic and one-valued throughout the plane (when cut from $+1$ to $-\infty$), this equation, proved when $|z-1|>2$, persists throughout the cut plane.

Next take the integral $I_2$, defined by the equation

$$I_2\equiv\int_{-1}^{1}(1-t)^{\alpha+\lambda-\gamma}(1+t)^{\gamma-\beta+\lambda-1}(z-t)^{-\alpha-\lambda}\,dt,$$

in which the path of integration passes above the point $t=z$ when $I(z)\geqslant 0$, and below it when $I(z)<0$. Then $I_2$ is analytic in each part of the plane when the plane is cut along the whole length of the real axis.

Deforming the contour in the manner indicated in Figure 1, we see that

$$I_2=I_1+\int_{1}^{(z\mp)}(1-t)^{\alpha+\lambda-\gamma}(1+t)^{\gamma-\beta+\lambda-1}(z-t)^{-\alpha-\lambda}\,dt,$$

where the path of integration starts from $t=1$ and returns to it after encircling the point $t=z$ negatively or positively according as $I(z)$ is positive or negative.

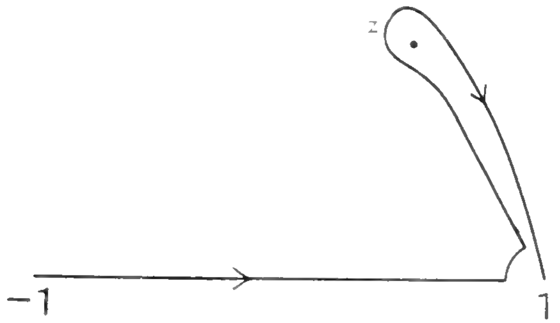



FIG. 1.

To evaluate this loop integral, take $|z-1|<2$ and write $1-t=(t-1)e^{\mp\pi i}$, so that

$$|\arg(t-1)|<\pi;$$

and then put
$$u=(t-1)/(z-1),$$

so that
$$t-1=u(z-1),\quad z-t=(1-u)(z-1),\quad t+1=2+u(z-1),$$

where $\arg(1-u)$ is zero on the first part of the loop and $\mp 2\pi$ on the second part.

Then

$$I_2 - I_1 = \int_0^{(1\mp)} \{u(z-1)\, e^{\mp\pi i}\}^{\alpha-\lambda-\gamma}\, 2^{\gamma-\beta+\lambda-1} \sum_{n=0}^{\infty} \frac{\Gamma(\beta-\lambda-\gamma+1+n)}{n!\,\Gamma(\beta-\lambda-\gamma+1)} \frac{u^n(1-z)^n}{2^n}$$

$$\times \{(1-u)(z-1)\}^{-\alpha-\lambda}(z-1)\,du$$

$$= \{1-e^{\pm 2\pi i(\alpha+\lambda)}\}\, 2^{\lambda-\beta}\, e^{\mp\pi i(\alpha+\lambda-\gamma)}\, \frac{\Gamma(\alpha+\lambda-\gamma+1)\Gamma(1-\alpha-\lambda)}{\Gamma(2-\gamma)} \left(\frac{z-1}{2}\right)^{1-\gamma}$$

$$\times F(\alpha+\lambda-\gamma+1,\ \beta-\lambda-\gamma+1;\ 2-\gamma;\ \tfrac{1}{2}-\tfrac{1}{2}z)$$

$$= \mp \frac{2\pi i e^{\pm\pi i\gamma} 2^{\lambda-\beta}\Gamma(\alpha+\lambda-\gamma+1)}{\Gamma(2-\gamma)\Gamma(\alpha+\lambda)} \left(\frac{z-1}{2}\right)^{1-\gamma} F(\alpha+\lambda-\gamma+1,\ \beta-\lambda-\gamma+1;\ 2-\gamma;\ \tfrac{1}{2}-\tfrac{1}{2}z).$$

Since it follows from the formulae of analytic continuation* that

$$I_1 = \frac{2^{\lambda-\beta}\,\Gamma(\gamma-\beta+\lambda)\,\Gamma(1-\gamma)}{\Gamma(1-\beta+\lambda)} F(\alpha+\lambda,\ \beta-\lambda;\ \gamma;\ \tfrac{1}{2}-\tfrac{1}{2}z)$$

$$+ \frac{2^{\lambda-\beta}\,\Gamma(\alpha+\lambda-\gamma+1)\,\Gamma(\gamma-1)}{\Gamma(\alpha+\lambda)} \left(\frac{z-1}{2}\right)^{1-\gamma} F(\alpha+\lambda-\gamma+1,\ \beta-\lambda-\gamma+1;\ 2-\gamma;\ \tfrac{1}{2}-\tfrac{1}{2}z),$$

it is easily shewn that

$$F(\alpha+\lambda,\ \beta-\lambda;\ \gamma;\ \tfrac{1}{2}-\tfrac{1}{2}z) = \frac{\Gamma(1-\beta+\lambda)\,\Gamma(\gamma)}{2^{\lambda-\beta+1}\pi i\,\Gamma(\gamma-\beta+\lambda)} \{\pm e^{\pm\pi i\gamma} I_1 \mp e^{\mp\pi i\gamma} I_2\};$$

strictly speaking, this result has not been proved when $\gamma$ is a positive integer, in view of the factor $\Gamma(1-\gamma)$ which occurs in the course of the reduction; but, since both sides of the final equation are analytic functions of $\gamma$ (except when $\gamma$ is zero or a negative integer), the equation holds also for the exceptional values of 1, 2, ... of $\gamma$, provided only that $|\lambda|$ is so large that the integrals $I_1$ and $I_2$ are convergent.

### 4. *The contours provided by the method of steepest descents.*

We now apply the principles of the method of steepest descents to the integral

$$\int (1-t)^{\alpha-\gamma}(1+t)^{\gamma-\beta-1}(z-t)^{-\alpha} \exp\left(-\lambda \log \frac{z-t}{1-t^2}\right) dt,$$

with a view to determining asymptotic expansions of $I_1$, $I_2$ when $|\lambda|$ is large.

The stationary points of $\log\{(z-t)/(1-t^2)\}$, *qua* function of $t$, are given by the equation

$$t^2 - 2zt + 1 = 0,$$

and so they are $\qquad t = e^{\zeta}, \quad t = e^{-\zeta}.$

The contours which are provided by the method of steepest descents are consequently arcs of the curves†

$$I \log \frac{z-t}{1-t^2} = I \log \frac{z-e^{\zeta}}{1-e^{2\zeta}}, \qquad I \log \frac{z-t}{1-t^2} = I \log \frac{z-e^{-\zeta}}{1-e^{-2\zeta}}.$$

Writing‡ $t = X + iY$, and supposing that $\phi_1$, $\phi_2$, $\phi_3$ are the angles which the vectors $t-1$, $t+1$, $t-z$ make with the $X$-axis, we see that each of the curves is such that $\phi_1 + \phi_2 - \phi_3$ is constant on it. These curves are portions of circular cubics.

* Barnes, *Proc. London Math. Soc.* (2) VI. (1908), p. 147.

† Riemann takes the real part of the logarithm constant and then applies the method of stationary phase; this procedure involves very great labour in obtaining terms of the asymptotic expansion following the dominant terms (cf. *Cambridge Philosophical Proceedings* XIX. (1917), p. 45).

‡ Of course $X$ and $Y$ are supposed to be real.

Since $\tan\left\{I\log\dfrac{t-z}{t^2-1}\right\}=\dfrac{-Y(X^2+Y^2+1)-\sinh\xi\sin\eta\,(X^2-Y^2-1)+2XY\cosh\xi\cos\eta}{X(X^2+Y^2-1)-\cosh\xi\cos\eta\,(X^2-Y^2-1)-2XY\sinh\xi\sin\eta}$,

the equations of the two cubics are

$$Y(X^2+Y^2+1)+\sinh\xi\sin\eta\,(X^2-Y^2-1)-2XY\cosh\xi\cos\eta$$
$$=\pm\tan\eta\,[X(X^2+Y^2-1)-\cosh\xi\cos\eta\,(X^2-Y^2-1)-2XY\sinh\xi\sin\eta],$$

and the cubics have nodes at $(e^{\xi}\cos\eta,\ e^{\xi}\sin\eta)$ and $(e^{-\xi}\cos\eta,\ -e^{-\xi}\sin\eta)$ respectively.

We shall consider fully (§§ 5–7) the properties of the cubic $(S_2)$ obtained by taking the upper sign in this equation, and deduce (§ 8) comparatively briefly the properties of the other cubic $(S_1)$; the simplest mode of passing from one cubic to the other is by changing the signs of $\xi$ and $\eta$ throughout the work.

### 5. *Properties of the cubic $S_2$.*

To express the coordinates of any point on $S_2$ as rational functions of a parameter $\mu$, we write the equation of the curve in the form

$$(Y\cos\eta-X\sin\eta)\{X(X-e^{\xi}\cos\eta)+Y(Y-e^{\xi}\sin\eta)\}-e^{-\xi}(X-e^{\xi}\cos\eta)(Y-e^{\xi}\sin\eta)=0;$$

and if we now put
$$Y-e^{\xi}\sin\eta=\mu\,(X-e^{\xi}\cos\eta),$$

we find after some straightforward algebra that

$$X=e^{\xi}\cos\eta+\frac{\mu^2e^{\xi}\cos\eta\sin\eta+\mu\,(e^{\xi}\cos2\eta-e^{-\xi})-e^{\xi}\cos\eta\sin\eta}{(\sin\eta-\mu\cos\eta)(1+\mu^2)},$$

$$Y=e^{\xi}\sin\eta+\frac{\mu\,\{\mu^2e^{\xi}\cos\eta\sin\eta+\mu\,(e^{\xi}\cos2\eta-e^{-\xi})-e^{\xi}\cos\eta\sin\eta\}}{(\sin\eta-\mu\cos\eta)(1+\mu^2)}.$$

The only real asymptote is $Y\cos\eta-X\sin\eta=e^{-\xi}\cos\eta\sin\eta$.

The curve degenerates (into a straight line and a circle) only when $\eta$ is $0,\ \pm\frac{1}{2},\ \pm\pi$, and, in each of these cases, the curves $C_1$ and $C_2$ coincide. It will, however, appear later that the degeneration when $\eta=\pm\frac{1}{2}\pi$ does not affect the analysis; but in the cases $\eta=0,\ \pm\pi$, the difficulty has to be surmounted by taking as contour not a portion of the degenerate cubic, but a slightly different curve (§ 7).

To return to the non-degenerate cubic, we notice that the effect of changing the sign of $\eta$ is to reflect the curve in the axis of $X$; while the effect of writing $\pm\pi-\eta$ for $\eta$ is to reflect the curve in the axis of $Y$. We can consequently derive the shape of the curve for any admissible value of $\eta$ by considering the shape of a curve of the family for which $\eta$ is a positive acute angle.

We now construct the following table of values of $t\,(\equiv X+iY)$ and $\mu$:

| | 1 | 2 | 3 | 4 | 5 | 6 |
|---|---|---|---|---|---|---|
| $t$ . . | $\infty$ | $z$ | $1$ | $e^{\xi}\cos\eta$ | $ie^{\xi}\sin\eta$ | $-1$ |
| $\mu$ . . | $\tan\eta$ | $\coth\xi\tan\eta$ | $\dfrac{e^{\xi}\sin\eta}{e^{\xi}\cos\eta-1}$ | $\infty$ | $0$ | $\dfrac{e^{\xi}\sin\eta}{e^{\xi}\cos\eta+1}$ |

The parameters and complex coordinates of these six special points will be denoted by attaching the suffixes 1, 2, ... 6 to $\mu$ and $t$.

It is easy to see from the table that (*when $\eta$ is a positive acute angle*) as $\mu$ increases from $\tan\eta$ to $+\infty$, and then from $-\infty$ to $\tan\eta$ again, it passes in succession through the values* $\mu_1$, $\mu_2$, ($\mu_3$ or $\mu_4$), $\mu_5$, $\mu_6$, $\mu_1$.

To determine the positions of the node relative to these six points, we observe that the parameters of the node are the roots of the quadratic $g(\mu)=0$, where

$$g(\mu)=\mu^2 e^\xi \cos\eta\sin\eta+\mu(e^\xi\cos 2\eta-e^{-\xi})-e^\xi\sin\eta\cos\eta;$$

now

$$g(\mu_4)>0>g(\mu_5),$$

and

$$g(\mu_3)=2e^{-\xi}\sin\eta(\cosh\xi-\cos\eta)/(e^\xi\cos\eta-1)^2>0,$$

$$g(\mu_1)=-e^{-\xi}\tan\eta<0,$$

$$g(\mu_2)=e^{-\xi}\sin\eta\cos\eta(\coth^2\xi\tan^2\eta+1)>0.$$

Hence one of the parameters of the node lies in the intervals $(\mu_3, \mu_5)$ and $(\mu_4, \mu_5)$, while the other lies in the interval $(\mu_1, \mu_2)$.

Hence, if a point starts from infinity and traverses the entire length of the cubic, it passes through the points of interest in the following order:

$$\infty,\quad e^\zeta\,(\text{node}),\quad z,\quad 1 \text{ or } e^\xi\cos\eta,\quad e^\zeta\,(\text{node}),\quad ie^\xi\sin\eta,\quad -1,\quad \infty$$

(the points being specified by their complex coordinates); if $\eta$ is a negative acute angle, the points are traversed in the same order; while if $\eta$ is obtuse (either positive or negative), the order is:

$$\infty,\quad e^\zeta\,(\text{node}),\quad z,\quad -1 \text{ or } e^\xi\cos\eta,\quad e^\zeta\,(\text{node}),\quad ie^\xi\sin\eta,\quad 1,\quad \infty.$$

The curve is shewn in Figures 2 and 3 in two cases, $\eta$ being a positive acute angle, and $e^\xi\cos\eta>1$ in 2 and $e^\xi\cos\eta<1$ in 3, while in Figure 4 $\eta=\frac{1}{2}\pi$. The portions of the curve from

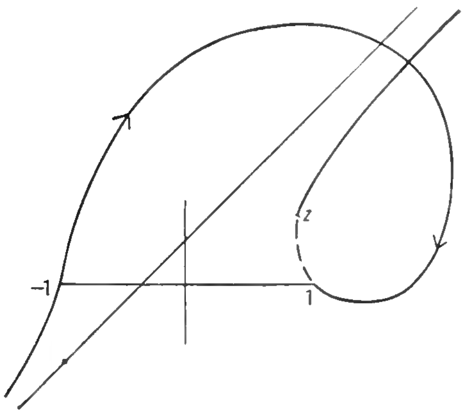



FIG. 2.

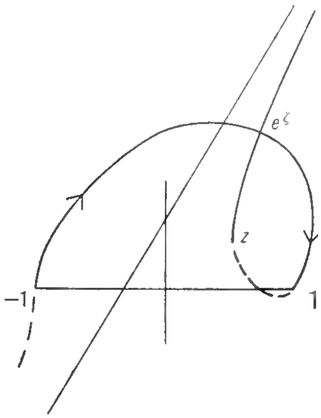



FIG. 3.

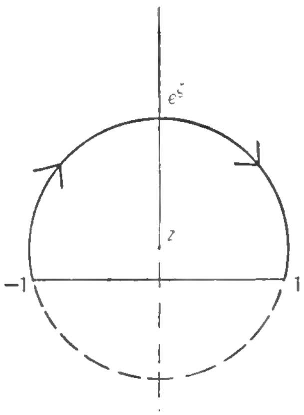



FIG. 4.

which the contour has to be selected are shewn in continuous lines; and it is obvious (even in the critical cases when $\eta=\pm\frac{1}{2}\pi$) that the cubic has an arc which passes from $-1$ to $1$ through the node without passing through† the critical points $z$ and $\infty$; and further, this arc cannot

* The sign of $e^\xi\cos\eta-1$ determines whether $\mu_3$ comes before or after $\mu_4$. When $e^\xi\cos\eta>1$, we prove that $\mu_2$ precedes $\mu_3$ by proving that $\coth\xi(e^\xi\cos\eta-1)<e^\xi\cos\eta$; this is true if $\operatorname{cosech}\xi\cos\eta<\coth\xi$, which is obviously the case.

† This is not true in the other critical cases $\eta=0, \pm\pi$; see the end of § 7.

be reconciled with the segment $(-1, 1)$ of the real axis without passing over the point $z$, and the arc is of the type specified for the second contour of § 3.

Further, as $t$ passes along the arc from $-1$ to $1$,

$$\tau \equiv \log\{(t-z)/(t^2-1)\} + \log(2e^{\zeta})$$

passes by a steady decrease from $+\infty$ to $0$ (at $t=e^{\zeta}$) and then increases steadily to $+\infty$ again.

Hence
$$I_2 = \left(\int_{\infty}^{0} + \int_{0}^{\infty}\right)(1-t)^{\alpha-\gamma}(1+t)^{\gamma-\beta-1}(z-t)^{-\alpha}\, 2^{\lambda} e^{\lambda\zeta} e^{-\lambda\tau} \frac{dt}{d\tau}\, d\tau.$$

## 6. *Transformations of the integrand of $I_2$.*

It is now possible to obtain the asymptotic expansion of $I_2$, by a consideration of the expansions of

$$(1-t)^{\alpha-\gamma}(1+t)^{\gamma-\beta-1}(z-t)^{-\alpha}(dt/d\tau),$$

in powers of $\tau$ on the two parts of the path near $t=e^{\zeta}$.

From the equation defining $\tau$, it is readily seen that

$$t - e^{\zeta} = \pm(1-e^{2\zeta-\tau})^{\frac{1}{2}}(1-e^{-\tau})^{\frac{1}{2}} - e^{\zeta}(1-e^{-\tau})$$
$$= \pm \sum_{s=0}^{\infty} a_s \tau^{s+\frac{1}{2}} + \sum_{s=0}^{\infty} b_s \tau^{s+1},$$

when $\tau$ is sufficiently small; if the upper sign refers to the arc joining $t=e^{\zeta}$ to $t=1$, then

$$a_0 = +(1-e^{2\zeta})^{\frac{1}{2}},$$

where the upper sign is given on the understanding that $a_0$ varies *continuously* with $\zeta$, so long as $I(e^{\zeta})$ does not change sign, and $a_0$ is positive when $\eta = \pm\frac{1}{2}\pi$, so that the saddle-point (*i.e.* the node of the cubic) is on the imaginary axis in the $t$-plane.

This result gives rise to an expansion of the form

$$(1-t)^{\alpha-\gamma}(1+t)^{\gamma-\beta-1}(z-t)^{-\alpha}\frac{dt}{d\tau} = \pm C \sum_{s=0}^{\infty} c_s \tau^{s-\frac{1}{2}} + \sum_{s=0}^{\infty} d_s \tau^{s},$$

where $c_0 = 1$ and
$$C = \tfrac{1}{2}(1-e^{\zeta})^{\alpha-\gamma+\frac{1}{2}}(1+e^{\zeta})^{\gamma-\beta-\frac{1}{2}}(z-e^{\zeta})^{-\alpha},$$

and the many-valued functions are specified by the conventions

$$|\arg(1-e^{\zeta})| < \pi, \quad |\arg(1+e^{\zeta})| < \pi,$$

since we took $\arg(1-t)$ and $\arg(1+t)$ to vanish when $-1 < t < 1$.

To determine $\arg(z-e^{\zeta}) \equiv \arg(-\sinh\zeta)$, we notice that it is to vary continuously with $\zeta$ so long as $I(z)$ does not change sign; and that, in the special cases when $\eta = \pm\frac{1}{2}\pi$, $z-e^{\zeta}$ is a pure imaginary whose sign is opposite to that of $\eta$; and that, as $t$ passes along the contour from $-1$ to $1$, $\arg(z-t)$ varies from $\arg(z+1)$ to $\arg(z-1) \mp 2\pi$; and hence that, in the special cases,

$$\arg(z-e^{\zeta}) = \mp\tfrac{1}{2}\pi.$$

Hence it follows that the equation

$$z - e^{\zeta} = \tfrac{1}{2}e^{-\zeta}(1-e^{\zeta})(1+e^{\zeta})$$

is always true in the sense that the argument of the expression on the left is the sum of the arguments of the factors on the right, and does not differ from the sum by a multiple of $2\pi$.

Therefore $$C = 2^{\alpha-1} e^{\alpha\zeta} (1-e^{\zeta})^{\frac{1}{2}-\gamma} (1+e^{\zeta})^{\gamma-\alpha-\beta-\frac{1}{2}}.$$

We shall next obtain a compact expression for the general coefficient $c_s$: it is evident that

$$Cc_s = \frac{1}{4\pi i}\int^{(0+,0+)} \left\{(1-t)^{\alpha-\gamma}(1+t)^{\gamma-\beta-1}(z-t)^{-\alpha}\frac{dt}{d\tau}\right\}\frac{d\tau}{\tau^{s+\frac{1}{2}}}$$
$$= \frac{1}{4\pi i}\int^{(\exp\zeta+)} \{(1-t)^{\alpha-\gamma}(1+t)^{\gamma-\beta-1}(z-t)^{-\alpha}\tau^{-s-\frac{1}{2}}(t-e^{\zeta})^{2s+1}\}\frac{dt}{(t-e^{\zeta})^{2s+1}},$$

the path of integration being a double circuit round the origin in the $\tau$-plane corresponding to a single circuit in the $t$-plane round $t = e^{\zeta}$.

Hence $Cc_s$ is half the coefficient of $(t-e^{\zeta})^{2s}$ in the expansion of

$$(1-t)^{\alpha-\gamma}(1+t)^{\gamma-\beta-1}(z-t)^{-\alpha}\tau^{-s-\frac{1}{2}}(t-e^{\zeta})^{2s+1}$$

in ascending powers of $t-e^{\zeta}$.

If we write $t = e^{\zeta}+T$, it is easy to shew that $c_s$ is equal to $(1-e^{2\zeta})^s$ multiplied by the coefficient of $T^{2s}$ in the expansion of

$$\left\{1-\frac{T}{1-e^{\zeta}}\right\}^{\alpha-\gamma}\left\{1+\frac{T}{1+e^{\zeta}}\right\}^{\gamma-\beta-1}\{1+T\operatorname{cosech}\zeta\}^{-\alpha}\left[\frac{1-e^{2\zeta}}{T^2}\log\left\{1+\frac{T^2}{1-e^{2\zeta}-2e^{\zeta}T-T^2}\right\}\right]^{-s-\frac{1}{2}}$$

in ascending powers of $T$.

The value of $c_0$ is unity, as has been already stated; and

$$c_1 = \tfrac{1}{2}(L+Me^{\zeta}+Ne^{2\zeta})/(1-e^{2\zeta}),$$

where
$$L = (\alpha+\beta-2\gamma+1)^2-\alpha+\beta-\tfrac{1}{2},$$
$$M = -2(\alpha+\beta-1)(\alpha+\beta-2\gamma+1),$$
$$N = (\alpha+\beta-1)^2+\alpha-\beta+\tfrac{1}{2}.$$

Next we consider the range of validity of the expansions $\pm C\Sigma c_s\tau^{s-\frac{1}{2}}+\Sigma d_s\tau^s$; the values $\pm 1$, $z$ of $t$ correspond to infinite values of $|\tau|$, and so the only finite singularities of the function of $\tau$ defined by the expansion are the points for which $dt/d\tau$ is zero or infinite (*i.e.* the points for which $t$ fails to be a monogenic function of $\tau$). These points are given by $t^2-2tz+1=0$, *i.e.* by $t = e^{\pm\zeta}$; and the corresponding values of $\tau$ are

$$\log\frac{2e^{\zeta}(e^{\mp\zeta}-z)}{e^{\pm 2\zeta}-1}.$$

These are the points $\tau = 2k\pi i$, $2\zeta+2k\pi i$, where $k$ is any integer. Hence, by Taylor's theorem, the radius of convergence of the expansion is $2\pi$ or $2|\zeta|$, whichever* is the smaller.

Further, it is easy to see that, throughout the $\tau$-plane (except in the immediate neighbourhood of the points $2k\pi i$, $2\zeta+2k\pi i$), we have

$$(1-t)^{\alpha-\gamma}(1+t)^{\gamma-\beta-1}(z-t)^{-\alpha}(dt/d\tau)-\left(\pm C\sum_{s=0}^{r}c_s\tau^{s-\frac{1}{2}}+\sum_{s=0}^{r}d_s\tau^s\right) = O(\tau^{r+\frac{1}{2}})+O(e^{K\tau}),$$

when $\tau$ is large, $K$ being a positive constant (*i.e.* independent of $\tau$) which depends on $\alpha$, $\beta$, $\gamma$, $z$, and $r$ being any fixed positive integer.

* Since $-\pi \leqslant \eta \leqslant \pi$, no one of the points $2\zeta+2k\pi i$ is nearer to the origin than $2\zeta$.

### 7. *The asymptotic expansion of $I_2$.*

It follows, by applying a general theorem* to the result of § 6, that we have the complete asymptotic expansion

$$\int_0^\infty (1-t)^{\alpha-\gamma}(1+t)^{\gamma-\beta-1}(z-t)^{-\alpha} e^{-\lambda\tau}(dt/d\tau)\,d\tau \sim \pm C\sum_{s=0}^{\infty} c_s \int_0^\infty \tau^{s-\frac{1}{2}} e^{-\lambda\tau}\,d\tau + \sum_{s=0}^{\infty} d_s \int_0^\infty \tau^s e^{-\lambda\tau}\,d\tau,$$

and so
$$I_2 \sim 2^{\lambda+1} C e^{\lambda\zeta} \sum_{s=0}^{\infty} c_s \Gamma(s+\tfrac{1}{2})\,\lambda^{-s-\frac{1}{2}},$$

the expansion being asymptotic in the sense of Poincaré when $|\lambda|$ is sufficiently large and $|\arg\lambda| \leqslant \frac{1}{2}\pi - \frac{1}{2}\delta$; provided that $\eta$ is not equal to $0$ or $\pm\pi$, as the integrand would then have a singularity for a positive value of $\tau$.

We shall now examine to what extent these restrictions can be removed.

*It can be shewn that the expansion is valid when $\eta = 0$ or $\pm\pi$*; for suppose that $\eta$ is slightly greater than $0$ (or $-\pi$); then instead of taking the contour to be the real axis in the $\tau$-plane, we take it to be† the ray $\arg\tau = -\frac{1}{2}\delta$; the modified integral is an analytic function of $\lambda$ when $-\frac{1}{2}\pi + \delta < \arg\lambda < \frac{1}{2}\pi$, and it is also an analytic function of $\zeta$ when $I(\zeta) \geqslant 0$ (or $-\pi$). Hence, making $\zeta$ assume the real value $\xi$ (or the value $\xi - \pi i$), we see that, when $\eta = +0$ or $-\pi + 0$, $I_2$ is equal to the modified integral, and also the asymptotic expansion is unaffected. To discuss the cases $\eta = -0$, $\pi - 0$ we proceed similarly, but we swing the contour round in the opposite direction; and we note that the expansion is the same whether $\eta = +0$ or $-0$.

Secondly, to extend the range of values of $\arg\lambda$, we observe that the process of swinging round the contour can be carried further, as shewn in Fig. 5. Take the two of the points $2\zeta + 2k\pi i$

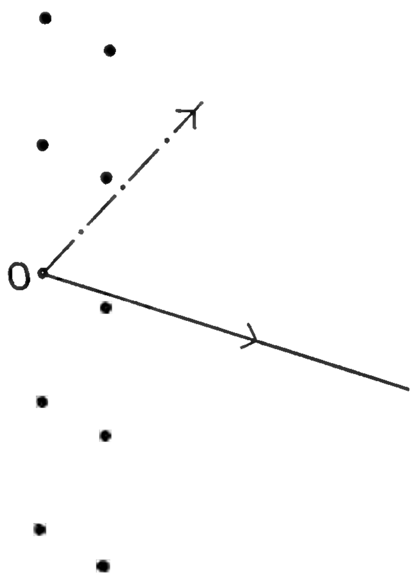



FIG. 5.

* *Proc. London Math. Soc.* (2) XVII. (1918), p. 133.

† The contour in the $t$-plane, corresponding to this deformed contour, is one which passes from $-1$ to $1$ and it is of a spiral form near each of these points.

which are nearest to the real axis (one on each side of it), and let the rays joining the origin to them be $\arg\tau = -\omega_1$, $\arg\tau = \omega_2$, so that $\omega_1$, $\omega_2$ are positive (or zero) acute angles.

When $\arg\lambda > 0$, we take the contour to be the ray $\arg\tau = -\omega_1 + \frac{1}{2}\delta$, and the modified integral (which has the asymptotic expansion already given) provides the analytic continuation of $I_2$ over the range for which $\arg\lambda < \frac{1}{2}\pi + \omega_1 - \delta$, provided that $|\lambda|$ is sufficiently large to make

$$K|\tau| - R(\lambda\tau) < 0.$$

Similarly, if $\arg\lambda < 0$, we swing the contour round to be the ray $\arg\tau = \omega_2 - \frac{1}{2}\delta$, and we get the asymptotic expansion of the analytic continuation of $I_2$ over the range $\arg\lambda > -\frac{1}{2}\pi - \omega_2 + \delta$.

*Hence, when $|\lambda|$ is sufficiently large, and $-\frac{1}{2}\pi - \omega_2 + \delta < \arg\lambda < \frac{1}{2}\pi + \omega_1 - \delta$, we have the asymptotic expansion of $I_2$ and its analytic continuation, given by the formula*

$$I_2 \sim 2^{\lambda+1} C e^{\lambda\zeta} \sum_{s=0}^{\infty} c_s \Gamma(s+\tfrac{1}{2}) \lambda^{-s-\frac{1}{2}}.$$

And the values of $\omega_1$, $\omega_2$ are given by the formulae

$$\omega_2 = \tan^{-1}(\eta/\xi), \quad -\omega_1 = \tan^{-1}\{(\eta - \pi)/\xi\},$$

when $\eta \geqslant 0$; and by the formulae

$$\omega_2 = \tan^{-1}\{(\eta + \pi)/\xi\}, \quad -\omega_1 = \tan^{-1}(\eta/\xi),$$

when $\eta \leqslant 0$, the symbol $\tan^{-1}$ in each case denoting an acute angle, positive or negative.

[The method obtaining the asymptotic expansion of $I_2$ when $\lambda$ does not lie in the specified sector of the plane is indicated in the first footnote to § 2.]

## 8. *The asymptotic expansion of $I_1$.*

We next consider the second contour $(S_1)$ of § 4, namely

$$I \log \frac{t - z}{t^2 - 1} = I \log \frac{e^{-\zeta} - z}{e^{-2\zeta} - 1}.$$

The analysis is derived from the preceding analysis by writing $(-\xi, -\eta)$ for $(\xi, \eta)$.

If $Y + e^{-\xi}\sin\eta = \mu(X - e^{-\xi}\cos\eta)$, we can construct (as in § 5) a table of corresponding values of $t\,(\equiv X + iY)$ and $\mu$.

| | 7 | 8 | 9 | 10 | 11 | 12 |
|---|---|---|---|---|---|---|
| $t =$ | $\infty$ | $-1$ | $-ie^{-\xi}\sin\eta$ | $1$ | $z$ | $e^{-\xi}\cos\eta$ |
| $\mu =$ | $-\tan\eta$ | $-\dfrac{e^{-\xi}\sin\eta}{1 + e^{-\xi}\cos\eta}$ | $0$ | $\dfrac{e^{-\xi}\sin\eta}{1 - e^{-\xi}\cos\eta}$ | $\coth\xi\tan\eta$ | $\infty$ |

We denote the parameters and complex coordinates of these six special points by attaching the suffices 7, 8, ... 12 to $\mu$ and $t$.

It is obvious from the table that, *when $\eta$ is a positive acute angle, $\mu_7, \mu_8, \ldots \mu_{12}$* are in ascending order of magnitude.

The parameters of the node are the roots of $h(\mu) = 0$, where

$$h(\mu) = \mu^2 e^{-\xi} \cos\eta \sin\eta - \mu(e^{-\xi}\cos 2\eta - e^{\xi}) - e^{-\xi}\cos\eta\sin\eta;$$

now
$$h(\mu_6) < 0, \quad h(\mu_{10}) = 2e^{-\xi}\sin\eta(\cosh\xi - \cos\eta)(1 - e^{-\xi}\cos\eta)^2 > 0,$$
$$h(\mu_{12}) > 0, \quad h(\mu_7) = -e^{\xi}\tan\eta < 0,$$

and so a point which traverses the entire length of the cubic passes through the points of interest in the following order:

$$\infty, \quad -1, \quad -ie^{\xi}\sin\eta, \quad e^{-\zeta}\,(\text{node}), \quad 1, \quad z, \quad e^{-\xi}\cos\eta, \quad e^{-\zeta}\,(\text{node}), \quad \infty.$$

Hence the arc of the cubic (Fig. 6) which passes from $-1$ to $1$ through the node is an admissible contour which lies entirely on one side of the real axis while the point $z$ lies on the other side, as shewn in Fig. 6.

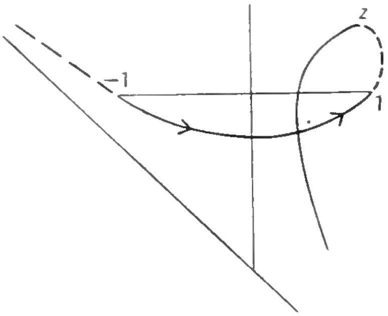



FIG. 6.

By suitable reflexions, we see that the cubic possesses these properties for all values of $\eta$ from $-\pi$ to $\pi$.

Hence, writing
$$\log\frac{t-z}{t^2-1} + \log 2e^{-\zeta} = \tau_1,$$

we get
$$I_1 = \left(\int_{\infty}^{0} + \int_{0}^{\infty}\right)(1-t)^{\alpha-\gamma}(1+t)^{\gamma-\beta-1}(z-t)^{-\alpha}\,2^{\lambda}e^{-\lambda\zeta}e^{-\lambda\tau_1}\frac{dt}{d\tau_1}\,d\tau_1.$$

Also, when $t$ is $e^{-\zeta}$, we have

$$1 - t = 1 - e^{-\zeta}, \quad 1 + t = 1 + e^{-\zeta}, \quad z - t = \tfrac{1}{2}e^{\zeta}(1 - e^{-\zeta})(1 + e^{-\zeta}),$$

where
$$|\arg(1 - e^{-\zeta})| \leqslant \pi, \quad |\arg(1 + e^{-\zeta})| \leqslant \pi.$$

We now proceed as in § 7, and we find that the asymptotic expansion of $I_1$, for large values of $\lambda$, is given by the formula

$$I_1 \sim 2^{\lambda+1}C' e^{-\lambda\zeta}\sum_{s=0}^{\infty} c_s'\,\Gamma(s+\tfrac{1}{2})\lambda^{-s-\frac{1}{2}},$$

where
$$C' = 2^{\alpha-1}e^{-\alpha\zeta}(1 - e^{-\zeta})^{\frac{1}{2}-\gamma}(1 + e^{-\zeta})^{\gamma-\alpha-\beta-\frac{1}{2}},$$

and the value of $c_0'$ is 1, while the general coefficient $c_s'$ is derived from $c_s$ by changing the sign of $\zeta$. In particular

$$c_1' = \tfrac{1}{2}(L + Me^{-\zeta} + Ne^{-2\zeta})(1 - e^{-2\zeta}),$$

where $L, M, N$ have the values given in § 6.

*There is, however, an important difference in the range of values of* $\arg\lambda$, *for which the asymptotic expansion of* $I_1$ *is valid.* For the singularities of the integrand are the points $\tau_1 = 2k\pi i,\ 2k\pi i - 2\zeta$, and the points $2k\pi i - 2\zeta$, being on the left of the imaginary axis, do not hamper the process of swinging round the path of integration; we may therefore swing it round so as to be either of the lines $\arg\tau_1 = \mp(\tfrac{1}{2}\pi - \tfrac{1}{2}\delta)$, according as $I(\lambda) \gtrless 0$: *and therefore the asymptotic expansion of* $I_1$ *is valid over the sector*

$$|\arg\lambda| \leqslant \pi - \delta,$$

provided that $|\lambda|$ is sufficiently large. For brevity we shall describe the sector for which

$$|\arg\lambda| \leqslant \pi - \delta,$$

as a *complete range* of values of $\arg\lambda$, while the sector for which $\arg\lambda$ lies between $-\tfrac{1}{2}\pi - \omega_2 + \delta$ and $\tfrac{1}{2}\pi + \omega_1 - \delta$ will be described as an *incomplete range* of values of $\arg\lambda$.

## 9. *Asymptotic expansions of hypergeometric functions.*

It is now possible to write down the asymptotic expansions of two independent solutions of the hypergeometric equation of the type under consideration; the formulae are as follows:

$$\left(\frac{z-1}{2}\right)^{-\alpha-\lambda} F\left(\alpha+\lambda,\ \alpha+\lambda-\gamma+1;\ \alpha-\beta+2\lambda+1;\ \frac{2}{1-z}\right)$$

$$\sim \frac{2^{\alpha+\beta}\Gamma(\alpha-\beta+2\lambda+1)}{\Gamma(\alpha+\lambda-\gamma+1)\Gamma(\gamma-\beta+\lambda)} e^{-(\alpha+\lambda)\zeta}(1-e^{-\zeta})^{\frac{1}{2}-\gamma}(1+e^{-\zeta})^{\gamma-\alpha-\beta-\frac{1}{2}} \sum_{s=0}^{\infty} c_s' \Gamma(s+\tfrac{1}{2})\lambda^{-s-\frac{1}{2}},$$

valid when $|\lambda|$ is large and $|\arg\lambda| \leqslant \pi - \delta$: and

$$F(\alpha+\lambda,\ \beta-\lambda;\ \gamma;\ \tfrac{1}{2}-\tfrac{1}{2}z) \sim \frac{\Gamma(1-\beta+\lambda)\Gamma(\gamma)}{\pi\Gamma(\gamma-\beta+\lambda)} 2^{\alpha+\beta-1}(1-e^{-\zeta})^{\frac{1}{2}-\gamma}(1+e^{-\zeta})^{\gamma-\alpha-\beta-\frac{1}{2}}$$

$$\times\left[e^{(\lambda-\beta)\zeta}\sum_{s=0}^{\infty} c_s\Gamma(s+\tfrac{1}{2})\lambda^{-s-\frac{1}{2}} + e^{\mp\pi i(\frac{1}{2}-\gamma)}e^{-(\lambda+\alpha)\zeta}\sum_{s=0}^{\infty} c_s'\Gamma(s+\tfrac{1}{2})\lambda^{-s-\frac{1}{2}}\right],$$

valid when $|\lambda|$ is large and $-\tfrac{1}{2}\pi - \omega_2 + \delta < \arg\lambda < \tfrac{1}{2}\pi + \omega_1 - \delta$.

In the second formula, the upper or lower sign is taken according as $I(z) \gtrless 0$, and, in deriving it, use has been made of the equation $1 - e^{\zeta} = e^{\zeta}(1-e^{-\zeta})e^{\mp\pi i}$.

The discontinuity in the second formula is only apparent; for if $z$ crosses the real axis between $\pm 1$, account has to be taken of the discontinuity in the value of $\eta$; while if $z$ crosses the real axis on the right of $z = 1$, we must have $|\arg\lambda| < \tfrac{1}{2}\pi - \delta$ for the crossing to be possible*, and then the second expansion is of exponentially lower order than the first.

* When $\eta = +0$ we must have

$$-\tfrac{1}{2}\pi - \tan^{-1}(\pi/\xi) + \delta < \arg\lambda < \tfrac{1}{2}\pi - \delta,$$

and, when $\eta = -0$, we must have

$$-\tfrac{1}{2}\pi + \delta < \arg\lambda < \tfrac{1}{2}\pi + \tan^{-1}(\pi/\xi) - \delta.$$

## PART II. ASYMPTOTIC EXPANSIONS OF LEGENDRE FUNCTIONS.

### 10. *The asymptotic expansion of $Q_n^m(z)$ when $|n|$ is large.*

As special cases of the formulae obtained in §§ 7–9 we can write down complete asymptotic expansions of the generalised Legendre functions $Q_n^m(z)$, $P_n^m(z)$ when $z$ is assigned and either $n$ or $m$ is large, $n$ and $m$ not being necessarily integers; the formulae agree with results previously obtained by Hobson and Barnes, but some of them are valid over a more extended range than has been hitherto assigned.

Let us first consider $Q_n^m(z)$ when $|n|$ is large and $|\arg n| \leqslant \pi - \delta$. We have*

$$\frac{2Q_n^m(z)\sin n\pi}{\sin(n+m)\pi} = \left(\frac{z+1}{z-1}\right)^{\frac{1}{2}m} \frac{\Gamma(n+1)\Gamma(n+m+1)}{\Gamma(2n+2)} \{\tfrac{1}{2}(z-1)\}^{-n-1} F\left(m+n+1, n+1; 2n+2; \frac{2}{1-z}\right),$$

when $|\arg(z \pm 1)| \leqslant \pi$, so that

$$\frac{z+1}{z-1} = \left(\frac{1+e^{-\zeta}}{1-e^{-\zeta}}\right)^2,$$

the arguments of both sides of this equation being equal and not differing by a multiple of $2\pi$.

But, with these conventions, if $\alpha = 1$, $\beta = 0$, $\gamma = 1 - m$, $\lambda = n$, we have

$$I_1 = \frac{2^n\,\Gamma(n+m+1)\,\Gamma(n-m+1)}{\Gamma(2n+2)} \{\tfrac{1}{2}(z-1)\}^{-n-1} F\left(m-n+1, n+1; 2n+2; \frac{2}{1-z}\right).$$

Therefore, by § 8, the asymptotic expansion of $Q_n^m(z)$ is given by the formula

$$Q_n^m(z) \sim \frac{\Gamma(n+1)}{\Gamma(n-m+1)} \frac{\sin(n+m)\pi}{\sin n\pi} \frac{(\pi/n)^{\frac{1}{2}} e^{-(n-1)\zeta}}{\sqrt{(1-e^{-2\zeta})}} \left\{1 + \sum_{s=1}^{\infty} \frac{c_s'\,\Gamma(s+\frac{1}{2})}{\Gamma(\frac{1}{2})\,n^s}\right\},$$

valid when $|\arg(z \pm 1)| \leqslant \pi$ and $|\arg n| \leqslant \pi - \delta$; and $c_1'$ is given by the formula

$$c_1' = \frac{8m^2 - 3 + e^{-2\zeta}}{4(1-e^{-2\zeta})} = m^2 - \tfrac{1}{2} + (m^2 - \tfrac{1}{4})\coth\zeta.$$

In the special case when $z = \cos\eta$, $Q_n^m(z)$ is defined† as

$$\tfrac{1}{2}\{Q_n^m(\cosh(0+i\eta)) + Q_n^m\cosh(0-i\eta)\},$$

where we may take $0 < \eta < \pi$.

If we write in turn $\zeta = 0 + i\eta$ and $\zeta = 0 - i\eta$, we get

$$e^{\zeta} = e^{\pm i\eta}\sqrt{(1-e^{-2\zeta})} = e^{\pm\frac{1}{2}i\eta \pm \frac{1}{4}\pi i}\sqrt{(2\sin\eta)},$$

since $\sqrt{(1-e^{-2\zeta})}$ is positive when $\zeta$ is a pure imaginary.

We thus obtain the complete asymptotic expansion

$$Q_n^m(\cos\eta) \sim \frac{\Gamma(n+1)}{\Gamma(n-m+1)} \frac{\sin(n+m)\pi}{\sin n\pi} \sqrt{\left(\frac{\pi}{2n\sin\eta}\right)} \cdot \left[\cos\{(\eta+\tfrac{1}{2})\eta + \tfrac{1}{4}\pi\} \cdot \left\{1 + \frac{m^2-\frac{1}{2}}{2n} + \dots\right\} - \frac{(m^2-\frac{1}{4})\cot\eta}{2n} \sin\{(n+\tfrac{1}{2})\eta + \tfrac{1}{4}\pi\} + \dots\right],$$

valid when $0 < \eta < \pi$, $|\arg n| < \pi - \delta$.

* This is in accordance with the definition given by Barnes, *Quarterly Journal* (*loc. cit.*), pp. 100, 107.

† This is also in accordance with Barnes' definition (p. 114). It differs from Hobson's definition, *Phil. Trans.* (*loc. cit.*), p. 471.

11. *The asymptotic expansion of $P_n^m(z)$ when $|n|$ is large.*

From the well-known formula*

$$P_n^m(z) = \frac{1}{\Gamma(1-m)}\left(\frac{z+1}{z-1}\right)^{\frac{1}{2}m} F(n+1, -n; 1-m; \tfrac{1}{2}-\tfrac{1}{2}z),$$

we at once derive the asymptotic expansion

$$P_n^m(z) \sim \frac{\Gamma(n+1)}{\Gamma(n-m+1)} \frac{e^{-\frac{1}{2}\zeta}}{(n\pi)^{\frac{1}{2}}(1-e^{-2\zeta})^{\frac{1}{2}}}$$

$$\times \left[ e^{(n+\frac{1}{2})\zeta} \sum_{s=0}^{\infty} \frac{c_s \Gamma(s+\frac{1}{2})}{\Gamma(\frac{1}{2}) n^s} + e^{\mp\pi i(m-\frac{1}{2})} e^{-(n+\frac{1}{2})\zeta} \sum_{s=0}^{\infty} \frac{c_s' \Gamma(s+\frac{1}{2})}{\Gamma(\frac{1}{2}) n^s} \right],$$

valid throughout the incomplete range of values of $\arg n$ given by

$$-\tfrac{1}{2}\pi - \omega_2 + \delta \leqslant \arg n \leqslant \tfrac{1}{2}\pi + \omega_1 + \delta.$$

In the special case in which $\eta = 0$ and $|\arg n| \leqslant \frac{1}{2}\pi - \delta$, the second expansion may be omitted and we have the simpler formula

$$P_n^m(\cosh \xi) \sim \frac{\Gamma(n+1)}{\Gamma(n-m+1)} \frac{e^{(n+\frac{1}{2})\xi}}{\sqrt{(2n\pi \sinh \xi)}} \sum_{s=0}^{\infty} \frac{c_s \Gamma(s+\frac{1}{2})}{\Gamma(\frac{1}{2}) n^s}.$$

To discuss the case in which $-1 < z < 1$, we write $z = \cos\eta$ where $0 < \eta < \pi$. Then, remembering that, with the usual definition,

$$P_n^m(\cos\eta) = e^{\frac{1}{2}m\pi i} P_n^m\{\cosh(0+\eta i)\},$$

we get
$$P_n^m(\cos\eta) \sim \frac{\Gamma(n+1)}{\Gamma(n-m+1)} \frac{1}{\sqrt{(2n\pi \sin\eta)}} \left[ e^{(\frac{1}{2}m\pi - \frac{1}{4}\pi + n\eta + \frac{1}{2}\eta)i} \sum_{s=0}^{\infty} \frac{c_s \Gamma(s+\frac{1}{2})}{\Gamma(\frac{1}{2} n)^s} \right.$$

$$\left. + e^{-(\frac{1}{2}m\pi - \frac{1}{4}\pi + n\eta + \frac{1}{2}\eta)i} \sum_{s=0}^{\infty} \frac{c_s' \Gamma(s+\frac{1}{2})}{\Gamma(\frac{1}{2}) n^s} \right];$$

this expansion is valid when $|\arg n| \leqslant \pi - \delta$.

The dominant terms are

$$P_n^m(\cos\eta) \sim \frac{\Gamma(n+1)}{\Gamma(n-m+1)} \left(\frac{2}{n\pi\sin\eta}\right)^{\frac{1}{2}} \left[ \cos\{(n+\tfrac{1}{2})\eta + \tfrac{1}{2}m\pi - \tfrac{1}{4}\pi\} + \frac{m^2-\frac{1}{4}}{2n}\cos\{(n+\tfrac{1}{2})\eta+\tfrac{1}{2}m\pi-\tfrac{1}{4}\pi\} \right.$$

$$\left. - \frac{m^2-\frac{1}{4}}{2n}\cot\eta \sin\{(n+\tfrac{1}{2})\eta + \tfrac{1}{2}m\pi - \tfrac{1}{4}\pi\} + \dots \right].$$

It is worthy of note that the asymptotic expansion of $P_n^m(\cos\eta)$ is valid over a more extended range of values of $\arg n$ than is given by Barnes' method (p. 155). The results of this section are otherwise equivalent to those given by him on pp. 152–161.

12. *Asymptotic expansions of Legendre functions whose order $m$ is large.*

These expansions need a rather more lengthy investigation than the expansions given in §§ 10, 11, since the cases $R(z) \geqslant 0$, $R(z) \leqslant 0$ have to be considered separately.

The formulae† which will be employed to obtain the expansions are

$$P_n^m(z) = \frac{2^m (z^2-1)^{-\frac{1}{2}m}}{\Gamma(1-m)} F(\tfrac{1}{2}n + \tfrac{1}{2} - \tfrac{1}{2}m, -\tfrac{1}{2}n - \tfrac{1}{2}m; 1-m; 1-z^2),$$

$$Q_n^m(z) = (z^2-1)^{-\frac{1}{2}(n+1)} \frac{2^{m-1}\Gamma(\frac{1}{2}n+\frac{1}{2}m+\frac{1}{2})\Gamma(\frac{1}{2}n+\frac{1}{2}m+1)}{\Gamma(n+\frac{3}{2})}$$

$$\times \frac{\sin(n+m)\pi}{\sin n\pi} F\left(\tfrac{1}{2}n+\tfrac{1}{2}m+\tfrac{1}{2}, \tfrac{1}{2}-\tfrac{1}{2}m+\tfrac{1}{2}; n+\tfrac{3}{2}; \frac{1}{1-z^2}\right),$$

which are valid when $R(z) \geqslant 0$.

* Hobson, p. 451; Barnes, p. 102.

† Barnes, pp. 120, 123.

If in the first of these formulae we write $-m$ for $m$, and then in the formulae of § 3 we write $\alpha=\beta=\frac{1}{2}n+\frac{1}{2}$, $\gamma=n+\frac{3}{2}$, $\lambda=\frac{1}{2}m$, we see that

$$P_n^{-m}(z)=\frac{2^{-\frac{3}{2}m-\frac{1}{2}n+\frac{1}{2}}(z^2-1)^{-\frac{1}{2}n-\frac{1}{2}}}{\Gamma(\frac{1}{2}m-\frac{1}{2}n)\,\Gamma(\frac{1}{2}m+\frac{1}{2}n+1)}\,I_1,$$

where, in $I_1$ we have to write $(1-z^2)^{-1}$ in place of $\frac{1}{2}(1-z)$ wherever $z$ occurs; so that $\zeta$ is now defined in terms of $z$ by the equation

$$\tfrac{1}{2}(1-\cosh\zeta)=(1-z^2)^{-1},$$

and therefore $$\cosh\zeta=\frac{z^2+1}{z^2-1},\qquad \sinh\zeta=\frac{\pm 2z}{z^2-1},$$

and, since $R(z)\geqslant 0$, the upper sign must be taken in order that we may have $|e^{-\zeta}|\leqslant 1$.

The asymptotic expansion of $P_n^{-m}(z)$ is therefore

$$P_n^{-m}(z)\sim\frac{2^{-m}}{\Gamma(\frac{1}{2}m-\frac{1}{2}n)\,\Gamma(\frac{1}{2}m+\frac{1}{2}n+1)}\left(\frac{z-1}{z+1}\right)^{\frac{1}{2}m}\left(\frac{2\pi}{m}\right)^{\frac{1}{2}}\left[1+\sum_{s=1}^{\infty}\frac{2^s c_s'\,\Gamma(s+\frac{1}{2})}{\Gamma(\frac{1}{2})\,m^s}\right],$$

where $c_s'(1+z)^{2s}(4z)^{-s}$ is the coefficient of $T^{2s}$ in the expansion of

$$\{1-\tfrac{1}{2}(1+z)T\}^{-\frac{1}{2}n-1}\{1+\tfrac{1}{2}z^{-1}(1+z)T\}^{\frac{1}{2}n}\{1-\tfrac{1}{2}z^{-1}(z^2-1)T\}^{-\frac{1}{2}n-\frac{1}{2}}$$
$$\times\left[\frac{(1+z)^2}{4zT^2}\log\left\{1+\frac{(1+z)^2T^2}{4z-2(z^2-1)T-(1+z)^2T^2}\right\}\right]^{-s-\frac{1}{2}};$$

and the asymptotic expansion is valid, when $R(z)\geqslant 0$, over a complete range (§ 8) of values of $\arg m$.

By using the second asymptotic expansion given in § 9, we find

$$Q_n^m(z)\sim 2^{m-1}\frac{\sin(n+m)\pi}{\sin n\pi}\,\frac{\Gamma(\frac{1}{2}m+\frac{1}{2}n+\frac{1}{2})\,\Gamma(\frac{1}{2}m-\frac{1}{2}n+\frac{1}{2})}{\sqrt{(2m\pi)}}\left[\left(\frac{z+1}{z-1}\right)^{\frac{1}{2}m}\left\{1+\sum_{s=1}^{\infty}\frac{2^s c_s\,\Gamma(s+\frac{1}{2})}{\Gamma(\frac{1}{2})\,m^s}\right\}\right.$$
$$\left.-e^{\mp n\pi i}\left(\frac{z-1}{z+1}\right)^{\frac{1}{2}m}\left\{1+\sum_{s=1}^{\infty}\frac{2^s c_s'\,\Gamma(s+\frac{1}{2})}{\Gamma(\frac{1}{2})\,m^s}\right\}\right];$$

this is valid, when $R(z)\geqslant 0$, over an incomplete range of values of $\arg m$; the upper or lower sign is taken according as $I\left(\frac{z+1}{z-1}\right)\lesseqgtr 0$, *i.e.* as $I(z)\gtreqless 0$, and $c_s$ is derived from $c_s'$ by changing the sign of $z$.

From the formula (Hobson, p. 462; Barnes, p. 109)

$$2Q_n^m(z)\,\Gamma(-m-n)\,\pi^{-2}\sin m\pi\sin n\pi=\frac{P_n^{-m}(z)}{\Gamma(1-m+n)}-\frac{P_n^m(z)}{\Gamma(1+m+n)},$$

we find $$P_n^m(z)\sim\frac{2^m\,\Gamma(\frac{1}{2}m+\frac{1}{2}n+\frac{1}{2})\,\Gamma(\frac{1}{2}m-\frac{1}{2}n+\frac{1}{2})}{\pi\sqrt{(2m\pi)}}\left[\sin m\pi\left(\frac{z+1}{z-1}\right)^{\frac{1}{2}m}\left\{1+\sum_{s=1}^{\infty}\frac{2^s c_s\,\Gamma(s+\frac{1}{2})}{\Gamma(\frac{1}{2})\,m^s}\right\}\right.$$
$$\left.-e^{\mp m\pi i}\sin n\pi\left(\frac{z-1}{z+1}\right)^{\frac{1}{2}m}\left\{1+\sum_{s=1}^{\infty}\frac{2^s c_s'\,\Gamma(s+\frac{1}{2})}{\Gamma(\frac{1}{2})\,m^s}\right\}\right].$$

This result is simplified by the disappearance of the first series in the special case when $m$ is a positive integer.

The general formula is valid, when $R(z)\geqslant 0$, over an incomplete range of values of $\arg m$; the special formula is true over a complete range.

Since $$Q_n^m(z)\,\Gamma(-m-n) = Q_n^{-m}(z)\,\Gamma(m-n)$$

(Barnes, p. 105), the asymptotic expansions when $R(z) \geqslant 0$ are completed.

We next consider the expansions when $R(z) \leqslant 0$.

From the formula (Hobson, p. 463; Barnes, p. 106)

$$P_n^m(-z) = P_n^m(z)\,e^{\mp n\pi i} - 2Q_n^m(z)\,\pi^{-1}\sin n\pi,$$

we see that, when $R(z) \geqslant 0$,

$$P_n^m(-z) \sim \frac{2^m\,\Gamma(\tfrac{1}{2}m+\tfrac{1}{2}n+\tfrac{1}{2})\,\Gamma(\tfrac{1}{2}m-\tfrac{1}{2}n+\tfrac{1}{2})}{\pi\sqrt{(2m\pi)}}\left[\sin m\pi\left(\frac{z-1}{z+1}\right)^{\frac{1}{2}m}\left\{1+\sum_{s=1}^{\infty}\frac{2^s c_s'\,\Gamma(s+\tfrac{1}{2})}{\Gamma(\tfrac{1}{2})\,m^s}\right\}\right.$$
$$\left.-\,e^{\pm m\pi i}\sin n\pi\left(\frac{z+1}{z-1}\right)^{\frac{1}{2}m}\left\{1+\sum_{s=1}^{\infty}\frac{2^s c_s\,\Gamma(s+\tfrac{1}{2})}{\Gamma(\tfrac{1}{2})\,m^s}\right\}\right].$$

Changing the sign of $z$, and noting that, when this is done, $c_s$ interchanges with $c_s'$ and $e^{m\pi i}$ with $e^{-m\pi i}$, we see that the expansion of $P_n^m(z)$ has the same form for all values of $z$ in an incomplete range of values of $\arg m$.

Next, taking the formula (Hobson, p. 463; Barnes, p. 106)

$$Q_n^m(-z) = -\,e^{\pm n\pi i}\,Q_n^m(z),$$

we obtain the asymptotic expansion of $Q_n^m(-z)$ when $R(z) \geqslant 0$; writing $-z$ for $z$, we see that, as in the case of $P_n^m(z)$, the asymptotic expansion of $Q_n^m(z)$ has the same form whether $R(z)$ be positive or negative; and from formulae already given connecting $P_n^{\pm m}$, $Q_n^{\pm m}$, it is evident that the same is true of $P_n^{-m}(z)$ and $Q_n^{-m}(z)$. The expansion of $P_n^{-m}(z)$ is valid for a complete range of values of $\arg m$ when $R(z) \geqslant 0$ only; the other expansions are valid for an incomplete range, as is that of $P_n^{-m}(z)$ when $R(z) \leqslant 0$.

The permanence of the form of the expansions when $R(z)$ passes through zero might have been anticipated* from the fact that, when $m$ is positive, $(z+1)^m$ and $(z-1)^m$ have the same modulus when $R(z)=0$, and that (as was pointed out by Stokes) discontinuities in asymptotic formulae usually occur in terms which are negligible in comparison with the dominant part of the expansion.

## PART III. MISCELLANEOUS PROPERTIES OF LEGENDRE FUNCTIONS.

### 13. *Definite integrals representing Legendre functions.*

We shall now obtain important and interesting definite integrals for $P_n(\cosh\zeta)$ and $Q_n(\cosh\zeta)$; they are derived from the formulae of §§ 3, 6, 8 by putting $\alpha=1$, $\beta=0$, $\gamma=1$, $\lambda=n$; this substitution gives

$$I_2 = \left(\int_{\infty}^{0} + \int_{0}^{\infty}\right)(\cosh\zeta - t)^{-1}\,2^n e^{n\zeta} e^{-n\tau}\,(dt/d\tau)\,d\tau,$$

where $$\tau = \log\frac{t-\cosh\zeta}{t^2-1} + \log 2e^{\zeta},$$

so that $$\frac{d\tau}{dt} = -\frac{(t^2-1)(t-\cosh\zeta)}{t^2-2t\cosh\zeta+1}.$$

* The permanence of the expansions may also be seen from the fact that $P_n^{-m}(z)$ is expressible in terms of $F(n+1,\ -n;\ m+1;\ \tfrac{1}{2}-\tfrac{1}{2}z)$, which is of the form described in Part IV below as type B.

Let $t_1, t_2$ be the values of $t$ corresponding to any assigned positive value of $\tau$, of which $t_1$ is on that part of the contour $S_2$ in the $t$-plane which joins the node to the point $t = +1$.

Then, since $t_1, t_2$ are the roots of the quadratic

$$t^2 - 1 = 2e^{\zeta-\tau}(t - \cosh\zeta),$$

we have

$$\cosh\zeta = (1 + t_1 t_2)/(t_1 + t_2),$$

and so

$$(\cosh\zeta - t_1)^{-1}(dt_1/d\tau) - (\cosh\zeta - t_2)^{-1}(dt_2/d\tau)$$

$$= \frac{t_1^2 - 1}{t_1^2 - 2t_1\cosh\zeta + 1} - \frac{t_2^2 - 1}{t_2^2 - 2t_2\cosh\zeta + 1}$$

$$= \frac{t_1 + t_2}{t_1 - t_2} - \frac{t_2 + t_1}{t_2 - t_1}$$

$$= 4e^{\zeta-\tau}/(t_1 - t_2).$$

Now

$$t_1, t_2 = e^{\zeta-\tau} \pm (1 - e^{-\tau})^{\frac{1}{2}}(1 - e^{2\zeta-\tau})^{\frac{1}{2}},$$

where the upper sign refers to $t_1$ if $\arg(1 - e^{2\zeta-\tau}) \to 0$ as $\tau \to \infty$.

Hence we have

$$I_2 = 2^{n+1}e^{(n+1)\zeta}\int_0^\infty \frac{e^{-(n+1)\tau}\,d\tau}{(1 - e^{-\tau})^{\frac{1}{2}}(1 - e^{2\zeta-\tau})^{\frac{1}{2}}},$$

provided* that $\eta$ is not zero or $\pm\pi$.

In like manner

$$I_1 = 2^{n+1}e^{-(n+1)\zeta}\int_0^\infty \frac{e^{-(n+1)\tau}\,d\tau}{(1 - e^{-\tau})^{\frac{1}{2}}(1 - e^{-2\zeta-\tau})^{\frac{1}{2}}},$$

provided* that $\eta$ is not $\pm\pi$.

It is desirable to modify these results slightly, by observing that

$$1 - e^{2\zeta-\tau} = e^{\zeta}\{-(1 + e^{-\tau})\sinh\zeta + (1 - e^{-\tau})\cosh\zeta\}$$

$$= e^{\mp\pi i}e^{\zeta}\{(1 + e^{-\tau})\sinh\zeta - (1 - e^{-\tau})\cosh\zeta\},$$

according as $I(\zeta) \gtrless 0$, where $\arg(\sinh\zeta)$ lies between $\pm\pi$.

Similarly

$$1 - e^{-2\zeta-\tau} = e^{-\zeta}\{(1 + e^{-\tau})\sinh\zeta + (1 - e^{-\tau})\cosh\zeta\};$$

and so we have the formulae

$$I_1 = 2^{n+1}e^{-(n+\frac{1}{2})\zeta}\int_0^\infty \frac{e^{-(n+1)\tau}}{(1 - e^{-\tau})^{\frac{1}{2}}}\{(1 + e^{-\tau})\sinh\zeta + (1 - e^{-\tau})\cosh\zeta\}^{-\frac{1}{2}}\,d\tau,$$

$$I_2 = 2^{n+1}e^{\pm\frac{1}{2}\pi i + (n+\frac{1}{2})\zeta}\int_0^\infty \frac{e^{-(n+1)\tau}}{(1 - e^{-\tau})^{\frac{1}{2}}}\{(1 + e^{-\tau})\sinh\zeta - (1 - e^{-\tau})\cosh\zeta\}^{-\frac{1}{2}}\,d\tau.$$

The first of these gives the definite integral for $Q_n(z)$, namely

$$Q_n(z) = e^{-(n+\frac{1}{2})\zeta}\int_0^\infty \frac{e^{-(n+1)\tau}}{(1 - e^{-\tau})^{\frac{1}{2}}}\{(1 + e^{-\tau})\sinh\zeta + (1 - e^{-\tau})\cosh\zeta\}^{-\frac{1}{2}}\,d\tau,$$

valid except when $z + 1$ is real and negative; when $-1 < z < 1$, the mode of approach to the real axis has to be specified.

* Allowance can be made for the exceptional cases by a suitable indentation of the contour at the points where $e^{\tau} = e^{\pm 2\zeta}$.

When $z$ is on the real axis, however, we define $Q_n(\cos\eta)$ as the mean of the values on either side of the cut; so that

$$Q_n(\cos\eta) = \tfrac{1}{2}Q_n\{\cosh(0+i\eta)\} + \tfrac{1}{2}Q_n\{\cosh(0-i\eta)\};$$

since $\sinh(0 \pm i\eta) = e^{\pm\frac{1}{2}\pi i}\sin\eta$ when $0<\eta<\pi$, we see that, *for values of* $\eta$ *between* 0 *and* $\pi$, *we have*

$$Q_n(\cos\eta) = \tfrac{1}{2}e^{(n+\frac{1}{2})i\eta+\frac{1}{4}\pi i}\int_0^\infty \frac{e^{-(n+1)\tau}}{(1-e^{-\tau})^{\frac{1}{2}}}\{(1+e^{-\tau})\sin\eta - i(1-e^{-\tau})\cos\eta\}^{-\frac{1}{2}}\,d\tau$$

$$+\tfrac{1}{2}e^{-(n+\frac{1}{2})i\eta-\frac{1}{4}\pi i}\int_0^\infty \frac{e^{-(n+1)\tau}}{(1-e^{-\tau})^{\frac{1}{2}}}\{(1+e^{-\tau})\sin\eta + i(1-e^{-\tau})\cos\eta\}^{-\frac{1}{2}}\,d\tau.$$

Similarly, from the formula

$$P_n(z) = \pm(I_2 - I_1)/(2^{n+1}\pi i),$$

we have

$$P_n(z) = \pi^{-1}e^{(n+\frac{1}{2})\zeta}\int_0^\infty \frac{e^{-(n+1)\tau}}{(1-e^{-\tau})^{\frac{1}{2}}}\{(1-e^{-\tau})\sinh\zeta - (1-e^{-\tau})\cosh\zeta\}^{-\frac{1}{2}}\,d\tau$$

$$+\pi^{-1}e^{-(n+\frac{1}{2})\zeta\pm\frac{1}{2}\pi i}\int_0^\infty \frac{e^{-(n+1)\tau}}{(1-e^{-\tau})^{\frac{1}{2}}}\{(1+e^{-\tau})\sinh\zeta + (1-e^{-\tau})\cosh\zeta\}^{-\frac{1}{2}}\,d\tau,$$

provided $\eta$ is not zero or $\pi$.

Making $\xi \to 0$, *we see that, if* $0<\eta<\pi$, *we have*

$$P_n(\cos\eta) = \pi^{-1}e^{(n+\frac{1}{2})i\eta-\frac{1}{4}\pi i}\int_0^\infty \frac{e^{-(n+1)\tau}}{(1-e^{-\tau})^{\frac{1}{2}}}\{(1+e^{-\tau})\sin\eta - i(1-e^{-\tau})\cos\eta\}^{-\frac{1}{2}}\,d\tau$$

$$+\pi^{-1}e^{-(n+\frac{1}{2})i\eta+\frac{1}{4}\pi i}\int_0^\infty \frac{e^{-(n+1)\tau}}{(1-e^{-\tau})^{\frac{1}{2}}}\{(1+e^{-\tau})\sin\eta + i(1-e^{-\tau})\cos\eta\}^{-\frac{1}{2}}\,d\tau.$$

These integrals are fundamental in the subsequent analysis.

## 14. *Some properties of the zeros of Legendre functions.*

We shall now shew how to obtain roughly the positions of the zeros of $P_n(\cos\theta)$ and $Q_n(\cos\theta)$.

When $0<\theta<\pi$, we see that, as $\tau$ increases from 0 to $\infty$,

$$\arg\{(1+e^{-\tau})\sin\theta - i(1-e^{-\tau})\cos\theta\}$$

varies monotonically from 0 to $\theta - \frac{1}{2}\pi$; and so

$$\arg\{(1+e^{-\tau})\sin\theta - i(1-e^{-\tau})\cos\theta\}^{-\frac{1}{2}}$$

varies monotonically from 0 to $\frac{1}{4}\pi - \frac{1}{2}\theta$.

Hence, since $e^{-(n+1)\tau}/(1-e^{-\tau})^{\frac{1}{2}}$ is *positive*, we see, by considering the definition of an integral as the limit of a sum, that the value of

$$\arg\int_0^\infty \frac{e^{-(n+1)\tau}}{(1-e^{-\tau})^{\frac{1}{2}}}\{(1+e^{-\tau})\sin\theta - i(1-e^{-\tau})\cos\theta\}^{-\frac{1}{2}}\,d\tau$$

lies between 0 and $\frac{1}{4}\pi - \frac{1}{2}\theta$.

We may consequently write

$$\int_0^\infty \frac{e^{-(n-1)\tau}}{(1-e^{-\tau})^{\frac{1}{2}}} \{(1+e^{-\tau})\sin\theta \mp i(1-e^{-\tau})\cos\theta\}^{-\frac{1}{2}} d\tau = \tfrac{1}{2}\pi W e^{\pm i\omega},$$

where $W$ is positive and $\omega$ is an angle (depending on $n$ and $\theta$) between 0 and $\frac{1}{4}\pi - \frac{1}{2}\theta$.

We thus obtain the formulae

$$P_n(\cos\theta) = W\cos\{(n+\tfrac{1}{2})\theta - \tfrac{1}{4}\pi + \omega\},\quad Q_n(\cos\theta) = \tfrac{1}{2}\pi W\cos\{(n+\tfrac{1}{2})\theta + \tfrac{1}{4}\pi + \omega\}.$$

Next it will be shewn that $(n+\frac{1}{2})\theta + \omega$ is an increasing function of $\theta$ when $n$ is fixed. We have

$$\frac{\pi P_n(\cos\theta)}{2Q_n(\cos\theta)} = \tan\{(n+\tfrac{1}{2})\theta + \tfrac{1}{4}\pi + \omega\},$$

and so

$$\frac{\partial}{\partial\theta}\{(n+\tfrac{1}{2})\theta + \tfrac{1}{4}\pi + \omega\} = 2\pi\left\{Q_n(\cos\theta)\frac{dP_n(\cos\theta)}{d\theta} - P_n(\cos\theta)\frac{dQ_n(\cos\theta)}{d\theta}\right\}\Big/\Big[\pi^2\{P_n(\cos\theta)\}^2 + 4\{Q_n(\cos\theta)\}^2\Big].$$

But, by applying a well-known theorem, due to Abel*, to Legendre's equation, we deduce that $\sin^2\theta\{Q_nP_n' - P_nQ_n'\}$ is constant; and, on writing $\theta = \frac{1}{2}\pi$ and making use of the values of $P_n(0)$, $P_n'(0)$, $Q_n(0)$, $Q_n'(0)$ given by Hobson, p. 469, and Barnes, pp. 121, 124, we get

$$\sin^2\theta\{Q_nP_n' - P_nQ_n'\} = -1.$$

Therefore

$$\frac{\partial}{\partial\theta}\{(n+\tfrac{1}{2})\theta + \tfrac{1}{4}\pi + \omega\} = 2/\{\pi\sin\theta W^2\} > 0,$$

which gives the result stated.

We can now obtain limits for the zeros of $P_n(\cos\theta)$; when $0 \leqslant \theta \leqslant \frac{1}{2}\pi$, we observe that $(n+\frac{1}{2})\theta - \frac{1}{4}\pi + \omega$ certainly lies between $(k-\frac{1}{2})\pi$ and $(k+\frac{1}{2})\pi$, $k$ being an integer, if

$$(k-\tfrac{1}{2})\pi < (n+\tfrac{1}{2})\theta - \tfrac{1}{4}\pi,\ (n+\tfrac{1}{2})\theta - \tfrac{1}{4}\pi + (\tfrac{1}{4}\pi - \tfrac{1}{2}\theta) \leqslant (k+\tfrac{1}{2})\pi,$$

*i.e.* if

$$\frac{(4k-1)\pi}{4n+2} \leqslant \theta \leqslant \frac{(2k+1)\pi}{2n}.$$

In this range of values of $\theta$, $\cos\{(n+\frac{1}{2})\theta - \frac{1}{4}\pi + \omega\}$ has the sign of $(-1)^k$; therefore, as $\theta$ increases from $(2k+1)\pi/(2n)$ to $(4k+3)\pi/(4n+2)$, $(n+\frac{1}{2})\theta - \frac{1}{4}\pi + \omega$ steadily increases from a value between $(k\pm\frac{1}{2})\pi$ to a value between $(k+1\pm\frac{1}{2})\pi$; hence its cosine changes sign once and only once. Thus the only zeros of $P_n(\cos\theta)$ in the range $0 \leqslant \theta \leqslant \frac{1}{2}\pi$ are in the intervals

$$[(2k+1)\pi/(2n),\ (4k+3)\pi/(4n+2)];$$

and there is one zero and only one in each of these intervals†.

When $\frac{1}{2}\pi \leqslant \theta \leqslant \pi$, $\omega$ is negative, so that the inequalities are replaced by

$$(k-\tfrac{1}{2})\pi \leqslant (n+\tfrac{1}{2})\theta - \tfrac{1}{2}\theta,\ (n+\tfrac{1}{2})\theta - \tfrac{1}{4}\pi \leqslant (k+\tfrac{1}{2})\pi;$$

and we get one zero and only one in each of the intervals

$$[(4k+3)\pi/(4n+2),\ (2k+1)\pi/(2n)]$$

and none outside these intervals.

The function $Q_n(\cos\theta)$ can be dealt with in a similar manner; we shall not give the details as the reader will have no difficulty in constructing the analysis.

* *Crelle* II. p. 22. See also Forsyth, *Differential Equations*, § 65. The dashes denote differentiations with regard to $\cos\theta$.

† None of these intervals can have the point $\frac{1}{2}\pi$ as an internal point *unless $n$ is an odd integer*, in which case the corresponding interval is evanescent, and we have the known result that $P_n(0) = 0$.

## Part IV. Asymptotic expansions of a system of hypergeometric functions.

### 15. *The system of hypergeometric functions with large constant elements.*

We shall now determine the asymptotic expansion of any hypergeometric function in which one or more of the constant elements is large, provided that, when more than one of the constants is large, the ratio of the large constants is approximately $\pm 1$. The Jacobi-Tchebychef functions discussed in Part I are the most obviously important functions of this nature, but others seem to be of sufficient interest to justify the very brief account which we shall give.

The functions which will be considered are of the form

$$F(\alpha+\epsilon_1\lambda,\ \beta+\epsilon_2\lambda;\ \gamma+\epsilon_3\lambda;\ x),$$

where $\alpha$, $\beta$, $\gamma$, $x$ are assigned, $|\lambda|$ is large and $\epsilon_1$, $\epsilon_2$, $\epsilon_3$ have the values $0, \pm 1$.

There are obviously 27 sets of values of $(\epsilon_1, \epsilon_2, \epsilon_3)$, but of course the set $(0, 0, 0)$ has to be omitted; and nine other sets may also be omitted on account of the symmetry of the hypergeometric function in its first two elements; thus $(1, -1, 1)$ is effectively equivalent to $(-1, 1, 1)$.

We shall take the hypergeometric equations associated with the surviving seventeen functions and obtain asymptotic expansions of a fundamental pair of solutions of each equation. It will appear that the equations fall into four distinct types, according to the values of $(\epsilon_1, \epsilon_2, \epsilon_3)$ shewn in the following table:

| Case | $\epsilon_1$ | $\epsilon_2$ | $\epsilon_3$ | Type |
|---|---|---|---|---|
| 1 | 1 | −1 | 0 | A |
| 2 | 1 | 1 | 0 | A |
| 3 | −1 | −1 | 0 | A |
| 4 | 0 | 0 | 1 | B |
| 5 | 0 | 0 | −1 | B |
| 6 | 0 | 1 | 0 | B |
| 7 | 0 | −1 | 0 | B |
| 8 | 0 | 1 | 1 | B |
| 9 | 0 | −1 | −1 | B |
| 10 | 1 | 1 | 1 | B |
| 11 | −1 | −1 | −1 | B |
| 12 | 0 | 1 | −1 | C |
| 13 | 0 | −1 | 1 | C |
| 14 | 1 | −1 | 1 | C |
| 15 | 1 | −1 | −1 | C |
| 16 | 1 | 1 | −1 | D |
| 17 | −1 | −1 | 1 | D |

### 16. *Hypergeometric functions of type* A.

The reader will have observed that the functions of type A are those already investigated in Part I of this paper, in view of the fact that the function of case 1 is $F(\alpha+\lambda, \beta-\lambda; \gamma; x)$. We shall merely give a table, indicating the nature of the order of magnitude of the constant elements in the twenty-four hypergeometric functions connected with the equation which is satisfied by $F(\alpha+\lambda, \beta-\lambda; \gamma; x)$. By expressing any one of these functions in terms of the two fundamental integrals $I_1$ and $I_2$ introduced in Part I, the asymptotic expansions of the twenty-four solutions are at once obtained for a range of values of arg $\lambda$ greater than the half-plane $|\arg\lambda| \leqslant \frac{1}{2}\pi$; for values of $\lambda$ outside this range, we put $\lambda = -\lambda_1$, and then we obtain the asymptotic expansion of the function under consideration in terms of $\lambda_1$.

The complete set of functions of type A is given in the following table, the numbering of the solutions being that adopted by Forsyth, *Differential Equations*, §§ 120–121; the first three columns in the table give the coefficients of $\lambda$ in the corresponding elements of the hypergeometric functions connected with the solutions shewn in the fourth column.

| Coefficients of $\lambda$ | | | Functions | Case |
|---|---|---|---|---|
| 1 | −1 | 0 | (I)—(VIII) | 1 |
| 1 | 1 | 0 | (XVII), (XIX), (XXI), (XXIV) | 2 |
| −1 | −1 | 0 | (XVIII), (XX), (XXII), (XXIII) | 3 |
| 1 | 1 | 2 | (IX), (XII), (XIII), (XV) | — |
| −1 | −1 | −2 | (X), (XI), (XIV), (XVI) | — |

The simplest method of procedure is to express the function to be investigated in terms of the two fundamental solutions (I) and (IX) by means of the formulae given by Barnes, *Proc. London Math. Soc.* (2) VI. pp. 141–177, and then to use the equations connecting these solutions with the integrals $I_1$ and $I_2$.

### 17. *Hypergeometric functions of type* B.

This type consists of the twenty-four functions associated with the equation for which solution (I) is the function $F(\alpha, \beta; \gamma+\lambda; x)$; the coefficients of $\lambda$ in the constant elements of the twenty-four functions are as follows:

| Coefficients of $\lambda$ | | | Functions | Case |
|---|---|---|---|---|
| 0 | 0 | 1 | (I), (VIII) | 4 |
| 0 | 0 | −1 | (IV), (V) | 5 |
| 0 | 1 | 0 | (XI), (XII), (XIII), (XIV) | 6 |
| 0 | −1 | 0 | (IX), (X), (XV), (XVI) | 7 |
| 0 | 1 | 1 | (XVII), (XVIII), (XXIII), (XXIV) | 8 |
| 0 | −1 | −1 | (XIX), (XX), (XXI), (XXII) | 9 |
| 1 | 1 | 1 | (II), (VII) | 10 |
| −1 | −1 | −1 | (III), (VI) | 11 |

The two solutions which will be regarded as fundamental are the solution (I), namely $F(\alpha, \beta; \gamma+\lambda; x)$, and (VIII), namely $x^{1-\gamma-\lambda}(1-x)^{\gamma+\lambda-\alpha-\beta}F(1-\alpha, 1-\beta; \gamma+\lambda-\alpha-\beta+1; 1-x)$; it is evidently sufficient (since both are included in case 4) to obtain the asymptotic expansion of one of these functions.

The reader will easily verify that these solutions form a fundamental system when $|\lambda|$ is large.

We now investigate the asymptotic expansion of $F(\alpha, \beta; \gamma+\lambda; x)$; it will be found that, in the case of this function, the saddle-point, which is usually characteristic of the method of steepest descents, does not appear in the analysis.

We take the integral* $I_3 = \int_0^1 t^{\beta-1}(1-t)^{\gamma+\lambda-\beta-1}(1-xt)^{-\alpha}\,dt$,

and we observe that (when $\lambda$ is positive), $(1-t)^{\lambda}$ decreases steadily from 1 to zero as $t$ describes the path of integration.

Accordingly, writing $1-t=e^{-\tau}$, we have

$$I_3 = \int_0^\infty \{(1-e^{-\tau})^{\beta-1}e^{-\tau(\gamma-\beta)}(1-x+xe^{-\tau})^{-\alpha}\}\,e^{-\lambda\tau}\,d\tau.$$

Now, when $\tau$ is sufficiently small, we have

$$(1-e^{-\tau})^{\beta-1}e^{-\tau(\gamma-\beta)}(1-x+xe^{-\tau})^{-\alpha} = \tau^{\beta-1}\sum_{s=0}^{\infty} k_s\tau^s,$$

where $k_0=1$. Hence, when $R(\beta)\geqslant 0$, we have

$$I_3 \sim \sum_{s=0}^{\infty} \Gamma(\beta+s)\,k_s\,\lambda^{\beta-s};$$

the singularities of the integrand are at the points $\tau = 2n\pi i$, $2n\pi i + \log(1-x^{-1})$, and so, as in § 8, the expansion is valid over a complete range of values of $\arg\lambda$ when $|1-x^{-1}|\leqslant 1$, and over a certain incomplete range (greater than a half-plane) when $|1-x^{-1}|>1$.

Hence $$F(\alpha, \beta; \gamma+\lambda; x) \sim \frac{\Gamma(\gamma+\lambda)}{\Gamma(\gamma+\lambda-\beta)\,\lambda^{\beta}} \sum_{s=0}^{\infty} \frac{k_s\,\Gamma(\beta+s)}{\Gamma(\beta)\,\lambda^s}.$$

When $R(\beta)<0$ this result may be obtained by taking a Hankel-Pochhammer contour $(\infty; 0+)$ in the $\tau$-plane in place of the real axis.

The expansion on the right may be obtained formally by taking each term of the series for $F(\alpha, \beta; \gamma+\lambda; x)$, expanding in descending powers of $\lambda$, arranging the sum in powers of $\lambda$, and multiplying by the expansion of $\Gamma(\gamma+\lambda-\beta)/\Gamma(\gamma+\lambda)$ in descending powers of $\lambda$.

## 18. *Hypergeometric functions of type* C.

This type consists of the twenty-four functions associated with the equation for which solution (I) is the function $F(\alpha, \beta+\lambda; \gamma-\lambda; x)$; the coefficients of $\lambda$ in the constant elements of the twenty-four functions are as follows:

* If $x>1$, an indentation has to be made at $t=1/x$ in the path of integration.

| Coefficients of λ | | | Functions | Case |
|---|---|---|---|---|
| 0 | 1 | − 1 | (I), (IX) | 12 |
| 0 | − 1 | 1 | (IV), (XI) | 13 |
| 1 | − 1 | 1 | (XIV), (XIX) | 14 |
| 1 | 1 | − 1 | (XV), (XVIII) | 15 |
| − 1 | − 2 | − 1 | (II), (XII) | — |
| 1 | 2 | 1 | (III), (X) | — |
| 0 | 1 | 2 | (V), (XXI) | — |
| 0 | − 1 | − 2 | (VIII), (XXIII) | — |
| 0 | 2 | 1 | (XVI), (XX) | — |
| 0 | − 2 | − 1 | (XIII), (XVII) | — |
| 1 | 2 | 2 | (VI), (XXII) | — |
| − 1 | − 2 | − 2 | (VII), (XXIV) | — |

The two solutions which will be regarded as fundamental are the solution (V), namely $F(\alpha, \beta+\lambda;\ \alpha+\beta-\gamma+2\lambda+1;\ 1-x)$, for all values of $x$, and* (IV), namely

$$x^{1-\gamma-\lambda}(1-x)^{\gamma-\alpha-\beta-2\lambda} F(1-\alpha,\ 1-\beta-\lambda;\ 2-\gamma+\lambda;\ x),$$

inside the circle $|x| < 1$.

We take the integral

$$\int (-t)^{\beta+\lambda-1}(1-t)^{\gamma-\beta-2\lambda-1}(1-xt)^{-\alpha}\, dt.$$

Writing the integral in the form

$$\int (-t)^{\beta-1}(1-t)^{\gamma-\beta-1}(1-xt)^{-\alpha} \exp\left\{\lambda \log \frac{-t}{(1-t)^2}\right\} dt,$$

we find that the contour to be taken is given by the equation

$$I(\tau) \equiv I \log\{-4t/(1-t^2)\} = 0.$$

This curve consists of the negative part of the real axis, together with the circle $|t| = 1$. The saddle-point is $t = -1$, and $\tau$ vanishes there.

We therefore consider the integral†

$$I_3 = \int_{-\infty}^{0} (-t)^{\beta-1}(1-t)^{\gamma-\beta-1}(1-xt)^{-\alpha} \exp\left\{\lambda \log \frac{-t}{(1-t)^2}\right\} dt.$$

It is easy to shew that

$$I_3 = \frac{\Gamma(\beta+\lambda)\,\Gamma(\alpha-\gamma+\lambda+1)}{\Gamma(\alpha+\beta-\gamma+2\lambda+1)} F(\alpha, \beta+\lambda;\ \alpha+\beta-\gamma+2\lambda+1;\ 1-x),$$

when the $x$-plane is cut along the negative real axis. If $g_s$ is the coefficient of $T^{2s}$ in the expansion of

$$2^{2s}(1-T)^{\beta-1}(1-\tfrac{1}{2}T)^{\gamma-\beta-1}\left(1-\frac{xT}{1+x}\right)^{-\alpha}\left\{\frac{4}{T^2}\log\left(1+\frac{T^2}{4-4T}\right)\right\}^{-s-\frac{1}{2}},$$

we find in the usual manner that

$$I_3 \sim 2^{\gamma-\beta-2\lambda}(1+x)^{-\alpha}(\pi/\lambda)^{\frac{1}{2}} \sum_{s=0}^{\infty} \frac{g_s\,\Gamma(s+\frac{1}{2})}{\Gamma(\frac{1}{2})\,\lambda^s}.$$

* In part of the plane (viz. $|x| > 1$) we take (XI), namely $x^{1-\gamma+\lambda}(1-x)^{\gamma-\alpha-\beta-2\lambda} F(1-\alpha, \gamma-\alpha-\lambda; 1-\alpha+\beta+\lambda; 1/x)$, instead of (IV).

† The integrand does not converge at $t=1$ when $|\arg\lambda| < \frac{1}{2}\pi$, and hence we do not take the circle as contour.

It is found that the only finite singularities of $t$ *qua* function of $\tau$ are at the points in the $\tau$-plane for which $t=-1$; these are the points $\tau=2s\pi i$. The finite singularities of the other function are at the points in the $\tau$-plane at which $t=1/x$; these are the points $\tau=2s\pi i-\log\frac{-4x}{(x-1)^2}\cdot 2s\pi i$.

One of these points is on the real axis if $|x|=1$ or if $x$ is negative; and one of the points approaches the origin if, and only if, $x\to-1$; hence the expansion holds for an incomplete range of values of $\arg\lambda$ except at $x=-1$; and it holds for a complete range of values of $\arg\lambda$ when $|4x|>|x-1|^2$.

The domain of the complex variable $x$ for which this inequality holds is shaded* in Figure 7.

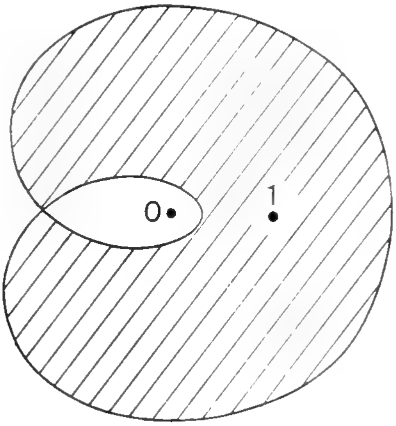



FIG. 7.

When $x=-1$ the expansion assumes a different form, since $(1-xt)^{-\alpha}$, when expanded in ascending powers of $\tau$, has its leading term $\tau^{-\frac{1}{2}\alpha}$ instead of $(1+x)^{-\alpha}$; it is easy to make the necessary modifications in the analysis.

Next take the integral

$$I_4=x^{1-\gamma-\lambda}(1-x)^{\gamma-\alpha-\beta-2\lambda}\int(-t)^{-\beta-\lambda}(1-t)^{\beta-\gamma+2\lambda}(1-xt)^{\alpha-1}\,dt.$$

The method of steepest descents gives the same potential contours as in the case of $I_3$, but now, in order to secure convergence, we take the circle $|t|=1$ (taken counter-clockwise starting from $t=1$) as contour; if $|x|<1$, by expanding in ascending powers of $x$, we get

$$I_4=-2\pi i\,x^{1-\gamma-\lambda}(1-x)^{\gamma-\alpha-\beta-2\lambda}\frac{\Gamma(1+\beta-\gamma+2\lambda)}{\Gamma(2-\gamma+\lambda)\,\Gamma(\beta+\lambda)}F(1-\alpha,\,1-\beta-\lambda;\,2-\gamma+\lambda;\,x).$$

If $|x|>1$, however, by expanding in descending powers of $x$, we get

$$\begin{aligned}I_4&=x^{\alpha-\gamma-\lambda}(1-x)^{\gamma-\alpha-\beta-2\lambda}\int(-t)^{\alpha-\beta-1-\lambda}(1-t)^{\beta-\gamma+2\lambda}\left(1-\frac{1}{xt}\right)^{\alpha-1}dt\\&=-2\pi i\,x^{\alpha-\gamma-\lambda}(1-x)^{\gamma-\alpha-\beta-2\lambda}\frac{\Gamma(1+\beta-\gamma+2\lambda)}{\Gamma(1+\alpha-\gamma+\lambda)\,\Gamma(1-\alpha+\beta+\lambda)}\\&\qquad\times F(1-\alpha,\,\gamma-\alpha-\lambda;\,1-\alpha+\beta+\lambda;\,1/x).\end{aligned}$$

* This is not a scale drawing but it affords a rough indication of the region in which the inequality holds. The small loop should be drawn very much smaller, as the points where the curve meets the real axis are $-1$, 0·172, 5·828.

To obtain the asymptotic expansion of $I_4$, we write

$$(1-t)^2(-4t)=e^{-\tau},$$

and then, if $g_s'$ is the coefficient of $T^{2s}$ in the expansion of

$$(-)^s 2^{2s}(1-T)^{\beta-1}(1-\tfrac{1}{2}T)^{\beta-\gamma}\left(1-\frac{xT}{1+x}\right)^{\alpha-1}\left\{\frac{4}{T^2}\log\left(1+\frac{T^2}{4-4T}\right)\right\}^{-s-\frac{1}{2}},$$

we get

$$I_4\sim -i2^{\beta-\gamma+1+2\lambda}(1+x)^{\alpha-1}x^{1-\gamma-\lambda}(1-x)^{\gamma-\alpha-\beta-2\lambda}(\pi/\lambda)^{\frac{1}{2}}\sum_{s=0}^{\infty}\frac{g_s'\Gamma(s+\frac{1}{2})}{\Gamma(\frac{1}{2})\lambda^s}.$$

The domains of values of $x$ in which this expansion holds for a complete range of values of $\arg\lambda$ are the unshaded portions of the plane in Fig. 7.

## 19. *Hypergeometric functions of type* D.

This type consists of the twenty-four functions associated with the equation for which solution (I) is $F(\alpha+\lambda, \beta+\lambda; \gamma+3\lambda; x)$: the coefficients of $\lambda$ in the constant elements of the twenty-four functions are as follows:

| Coefficients of λ | | | Functions | Case |
|---|---|---|---|---|
| 1 | 1 | 3 | (I) | — |
| 2 | 2 | 3 | (II) | — |
| −2 | −2 | −3 | (III) | — |
| −1 | −1 | −3 | (IV) | — |
| 1 | 1 | −1 | (V) | 16 |
| −2 | −2 | −1 | (VI) | — |
| 2 | 2 | 1 | (VII) | — |
| 1 | −1 | 1 | (VIII) | 17 |
| 1 | −2 | 0 | (IX), (X) | — |
| 1 | 2 | 0 | (XI), (XII) | — |
| 1 | 2 | 0 | (XIII), (XIV) | — |
| −1 | −2 | 0 | (XV), (XVI) | — |
| 1 | 2 | 3 | (XVII), (XVIII) | — |
| 1 | −2 | −3 | (XIX), (XX) | — |
| 1 | 2 | 1 | (XXI), (XXII) | — |
| 1 | 2 | 1 | (XXIII), (XXIV) | — |

The solutions which will be taken as fundamental are (I), viz. $F(\alpha+\lambda, \beta+\lambda; \gamma+3\lambda; x)$, and either (X), viz. $x^{-\beta-\lambda}F(\beta+\lambda, \beta-\gamma-2\lambda+1; \beta-\alpha+1; 1/x)$,

or (XXIV), viz.

$$x^{\beta-\gamma-2\lambda}(1-x)^{\gamma-\alpha-\beta+\lambda}F\{1-\beta-\lambda, \gamma-\beta+2\lambda; \gamma-\alpha-\beta+\lambda+1; (x-1)/x\}.$$

The analysis is somewhat similar to that employed in Part I; we take the integrals*

$$I_5, I_6=\int_0^1 t^{\beta+\lambda-1}(1-t)^{\gamma-\beta-2\lambda-1}(1-xt)^{-\alpha-\lambda}\,dt.$$

* It is supposed that, near $t=0$, $|\arg t|<\pi$ and $\arg(1-t)$ and $\arg(1-xt)$ are small.

The path for $I_5$ is reconcilable with the real axis without crossing over the singularity $t_3 \equiv 1/x$: the path for $I_6$ passes above or below this point according as the point is above or below the real axis.

It is readily verified that

$$I_5 = \frac{\Gamma(\gamma - \beta + 2\lambda)\,\Gamma(\beta + \lambda)}{\Gamma(\gamma + 3\lambda)} F(\alpha + \lambda, \beta + \lambda; \gamma + 3\lambda; x),$$

$$I_6 = \pm 2\pi i e^{\mp\pi i(\alpha+\lambda)} x^{-\beta-\lambda} \frac{\Gamma(\beta + \lambda)}{\Gamma(\alpha + \lambda)\,\Gamma(1 - \alpha + \beta)}$$

$$\times F(\beta + \lambda, \beta - \gamma - 2\lambda + 1; \beta - \alpha + 1; 1/x) + e^{\mp 2\pi i(\alpha-\lambda)} I_5,$$

where the upper or lower sign is taken according as $I(x) \gtrless 0$, and it is supposed that $|\arg x| \leqslant \pi$.

In order to employ the method of steepest descents we have to determine the stationary points of $(1 - xt)\,t^{-1}(1 - t)^{-2}$ *qua* function of $t$; they are given by the roots of the quadratic

$$2xt^2 - 3t + 1 = 0.$$

Put $9 - 8x = z$, it being understood that $|\arg z| < \pi$ [the cut from $x = 1$ to $x = +\infty$ in the $x$-plane insures this inequality being satisfied if we define $\arg(9 - 8x)$ to be zero when $x = 0$], and the stationary points are

$$t_1 \equiv 2/(3 + \sqrt{z}), \qquad t_2 = 2/(3 - \sqrt{z}).$$

The values of $(1 - xt)^{-1}(1 - t)^{-2}$ at $t_1$, $t_2$ are respectively

$$\tfrac{1}{8}(\sqrt{z} + 3)^3/(\sqrt{z} + 1), \qquad \tfrac{1}{8}(\sqrt{z} - 3)^3/(\sqrt{z} - 1).$$

We shall now discuss, by electrical methods*, the topography of the contours in the $t$-plane (for all assigned values of $z$) which are supplied by the application of the method of steepest descents.

If we write

$$\frac{1 - xt}{t(1 - t)^2} = e^{V + iW},$$

where $V$ and $W$ are real, it is evident that $V$ is the potential at the point $t$ due to a two-dimensional electrical distribution consisting of line charges through the points $0$, $1$, $1/x$ in the $t$-plane, the charges per unit length of the lines being in the ratios $1 : 2 : -1$.

The points $t_1$, $t_2$ are the only points of equilibrium; and the curves on which $V$ and $W$ respectively are constant are the equipotentials and the lines of force.

By straightforward algebra it is seen that the equipotentials are bicircular sextics and the lines of force are portions of circular quartics; the quartics pass through the points $t = 0, 1, 1/x$, and have a node at $t = 1$ and two real perpendicular asymptotes; the curves on which $W$ is constant consist of portions of the quartics with end points at the points $0, 1, 1/x, \infty$; and it is these portions of quartics (ending at the points $0$, $1$ only, when $R(\lambda) > 0$, to secure the convergence of $I_5$ and $I_6$) which are required by the method of steepest descents. We can now consider the topography of the different equipotentials obtained by varying $V$ from $+\infty$ to $-\infty$.

* I should have preferred to have employed the algebraic methods of Part I in discussing the forms of the contours instead of this combination of geometrical and electrical theories; but as the contours are portions of quartics with (in general) only two nodes, they are not unicursal curves, and so the algebra appeared intractible. The investigation actually given is, I think, quite rigorous.

When $V$ is large and positive, the equipotential consists of two small ovals* surrounding the positive charges at 0, 1 respectively. As $V$ decreases, the ovals increase in size, until we reach an equipotential through that one of the equilibrium points at which the potential is higher: this equipotential has a node which may arise in one or other of two ways, (I) by the two ovals uniting to form a figure-of-eight or (II) by previously distinct parts of the *same* oval bending round towards one another and uniting. Case (I) is shewn in Fig. 8 and case† (II) in Fig. 9.

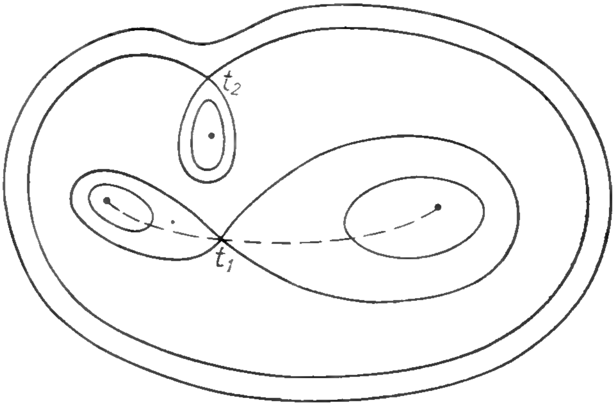



FIG. 8.

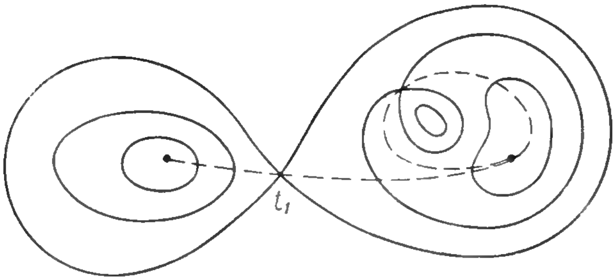



FIG. 9.

First take case (I). As $V$ decreases further the equipotential becomes a single oval surrounding the figure-of-eight and this form persists until we reach the equipotential through the equilibrium point with lower potential; the node at the equilibrium point can only be formed by distinct parts of the oval uniting to surround a portion of the plane not previously enclosed; this area having a portion of an equipotential as its complete boundary must contain a charge; this can only be the charge at $t_3 \equiv 1/x$. Subsequent equipotentials consist of two ovals, a large one surrounding all three charges and a small one inside the former surrounding the charge at $t_3$ only.

* The word oval is used, in the absence of a more suitable term, to mean a closed branch of a curve without nodes or cusps; it is not supposed that the branch has no inflections.

† It might have been anticipated that the left-hand oval could join up with itself; but, as will be seen later, this phenomenon does not occur when the charges have the proportions of those under discussion. Figs. 8 and 9 are not drawn to scale, but merely indicate the general topography of the plane; the dotted curves are lines of force.

Next we take case (II). As $V$ decreases further, the equipotentials become tripartite, consisting of an oval round the charge at 0, another round the charges at 1 and $t_3$, and a third, inside the second, round the charge at $t_3$ only. This form persists until the first two unite at the remaining equilibrium-point to form a figure-of-eight: and subsequent equipotentials are bipartite, consisting of a large oval round all three charges and a smaller one round the charge at $t_3$ only.

If now we regard $x$ (and therefore $z$) as a variable, the transition from case (I) to case (II) can only occur when $z$ passes through a value which makes the nodal equipotentials coincident; *i.e.* when $z$ satisfies the equation

$$\left|\frac{(\sqrt{z}+3)^3}{\sqrt{z}+1}\right| = \left|\frac{(\sqrt{z}-3)^3}{\sqrt{z}-1}\right|.$$

The curve in the $x$-plane on which this equation is satisfied is shewn* in Fig. 10; the simplest form of the equation of the curve is obtained by writing $z \equiv re^{i\theta}$ $(r \geqslant 0,\ -\pi \leqslant \theta \leqslant \pi)$, when the equation of the curve reduces to

$$r = 6\sqrt{3}\cos\tfrac{1}{2}\theta - 9,$$

together with the coincident rays $\cos\frac{1}{2}\theta = 0$ (*i.e.* $\theta = \pm\pi$).

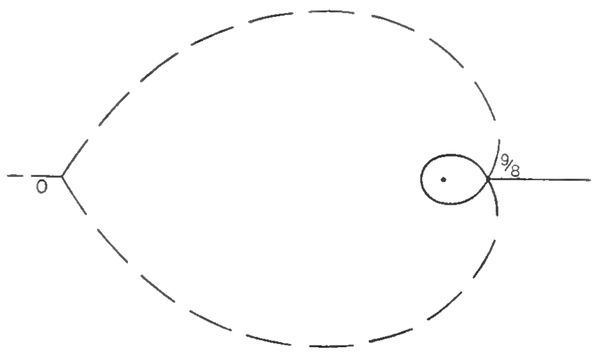



FIG. 10.

It is easy to see that when $x$ is *outside* the curve of Fig. 10 and fairly near the origin (so that $\sqrt{z}$ is comparatively nearly equal to 3), the potential at $t_1$ is higher than that at $t_2$. And, when $|x|$ is very small the charge at $t_3 (\equiv 1/x)$ has little influence on the form of the equipotentials moderately near 0 and 1; and so the equipotentials moderately near 0 and 1 have nearly the form which they would have if the only existing charges were at 0 and 1. *Hence, when $x$ is outside the curve of Fig.* 10, *the equipotentials have the configuration of case* (I); and the node of the figure-of-eight is at $t_1$, while the node of the other nodal equipotential (which may be described as a closed crescent) is at $t_2$.

When $|x-1|$ is small, so are $|t_3-1|$ and $|z-1|$; and, if we consider the special case in which $t_3 - 1$ is positive†, the equilibrium point $t_2$ is on the right of $t_3$ and the potential there, viz. $\log\{\frac{1}{8}(3-\sqrt{z})^3/(\sqrt{z}-1)\}$, is much higher than the potential at $t_1$, and $t_1$ is near the point $\frac{1}{2}$; hence we have the configuration of case (II) as shewn in Fig. 9; and so we have the configuration

* Fig. 10 is drawn to scale: the dotted curve is $r = 9 - 6\sqrt{3}\,|\sin\frac{1}{2}\theta|$, which will be required subsequently.

† So that $x$ is just less than 1, and $z$ is just greater than 1.

of Fig. 9 whenever $x$ is *inside** the curve of Fig. 10. It is to be noted that $t_1$ is *always* the node of the figure-of-eight.

[As a confirmation of these results, take $|x|$ large, when the equilibrium points are both near the origin; when we take $z$ positive, $t_1$ is nearer the origin than $t_2$, and the potential of $t_1$ is higher than that of $t_2$; thus we obtain the configuration of case (I). If now we vary arg $z$ from 0 to $\pm\pi$, keeping $|z|$ fixed, $t_3$ describes an arc round the origin, driving the node of the figure-of-eight round the origin in front of it with approximately half its own angular velocity. When arg $z$ becomes $\pm\pi$, the two parts of the figure-of-eight unite behind $t_3$, and we get a degenerate equipotential with two nodes, near the imaginary axis, which are symmetrically placed with respect to the real axis.]

*We shall now shew that part of the line of force through $t_1$ always passes from 0 to 1 and is reconcilable with the real axis without crossing over $t_3$.*

First take the configuration of case (I). The line of force through $t_1$ has a node there, and *one* of its branches (in the neighbourhood of $t_1$) lies *inside* the loops of the eight. This branch of the line of force cannot emerge from the loops of the eight before passing through an equilibrium point, and no such point exists. The line of force therefore terminates at 0 and 1. Further, when $|x|$ is small† the line of force nearly coincides with the real axis and $t_3$ is at a great distance from the origin, while $t_1$ and $t_3$ are on the opposite sides of the real axis. Therefore the line of force is reconcilable with the real axis. Now vary $x$, and we see that the line of force remains reconcilable with the real axis so long as $t_3$ does not cross the line of force or the real axis. But as $x$ varies, $t_3$ *cannot* cross the line of force without entering the figure of eight, *i.e.* without $x$ crossing the curve of Fig. 10; and since the variation in $x$ may be supposed to take place without $x$ crossing the real axis (since the initial value of $x$, with $|x|$ small, may have a positive or negative imaginary part, as we please), we see that whatever be the position of $x$ (outside the curve of Fig. 10), a line of force passes from 0 to 1 through $t_1$, and this line of force, so far as the point $t_3$ is concerned, is reconcilable with the real axis.

Next take the configuration of case (II). In the special case when $1-x$ is positive, $t_3$ is on the right of 1, and one branch of the line of force through $t_1$ passes straight from 0 to 1.

As we vary $x$ continuously the form of the line of force through $t_1$ varies continuously except when $x$ passes through such a value that the line of force has another node: but the quartic of which a line of force forms part can only have two nodes (other than the point $t=1$) if

$$\arg\frac{(\sqrt{z}+3)^3}{\sqrt{z}+1} - \arg\frac{(\sqrt{z}-3)^3}{\sqrt{z}-1}$$

is zero or an *even* multiple of $\pi$.

Now this difference is a multiple of $\pi$ only when

$$r = 9 \pm 6\sqrt{3}\sin\tfrac{1}{2}\theta,$$

or when $\sin\frac{1}{2}\theta = 0$, where, as previously, $z = re^{i\theta}$ $(r>0, -\pi \leqslant \theta \leqslant \pi)$; the difference is an odd multiple of $\pi$ on the branch $r = 9 + 6\sqrt{3}\,|\sin\frac{1}{2}\theta|$ near $z=9$, and so the difference is an even multiple of $\pi$ only when $\sin\frac{1}{2}\theta = 0$ or $r = 9 - 6\sqrt{3}\,|\sin\frac{1}{2}\theta|$; this curve is shewn by a dotted line in Fig. 10, and it lies wholly *outside* the curve $r = 6\sqrt{3}\cos\frac{1}{2}\theta - 9$. It follows that, as $x$ varies inside the continuous curve of Fig. 10, the general configuration of the branch of the line of force from 0 to 1 through $t_1$ does not change, but always lies inside the figure-of-eight and

* And so no case arises in which the closed-crescent equipotential contains the charges at 0, $1/x$ only; this justifies the statement made in footnote †, p. 304.

† Whether $I(x)$ be positive or negative.

does not go near* $t_3$. Also, as we may suppose that $t_1$ and $t_3$ do not cross the real axis, and as the branch of the line of force is reconcilable with the real axis, so far as $t_3$ is concerned, when $I(1-x)$ is very small and $R(1-x)$ is small and positive, it follows that, for all positions of $x$ inside the continuous curve of Fig. 10, the branch of the line of force under consideration is reconcilable with the real axis.

It is now a simple matter to apply the method of steepest descents to obtain the asymptotic expansion of $I_5$.

Writing $$(1-xt)\,t^{-1}(1-t)^{-2} = \tfrac{1}{8}\{(\sqrt{z}+3)^3(\sqrt{z}+1)\}\,e^{\tau},$$
where $\tau$ is positive, we get
$$I_5 = \left(\int_{\infty}^{0} + \int_{0}^{\infty}\right) t^{\beta-1}(1-t)^{\gamma-\beta-1}(1-xt)^{-\alpha}\left\{\frac{8(\sqrt{z}+1)}{(\sqrt{z}+3)^3}\right\}^{\lambda} e^{-\lambda\tau}\frac{dt}{d\tau}\,d\tau.$$
and $dt/d\tau$ is expansible (near $\tau=0$) in a series of ascending powers of $\tau^{\frac{1}{2}}$ commencing with a term in $\tau^{-\frac{1}{2}}$ whose coefficient is
$$\pm\tfrac{1}{2}z^{-\frac{1}{4}}(\sqrt{z}+1)/(\sqrt{z}+3);$$
and hence, after the manner of Part I, we obtain an asymptotic expansion for $I_5$, in descending powers of $\lambda$, of which the dominant term is given by the formula
$$I_5 \sim 2^{2\alpha+\beta-1}z^{-\frac{1}{4}}(\sqrt{z}+1)^{\gamma-\alpha-\beta}(\sqrt{z}+3)^{1-\gamma}\left\{\frac{8(\sqrt{z}+1)}{(\sqrt{z}+3)^3}\right\}^{\lambda}\left(\frac{\pi}{\lambda}\right)^{\frac{1}{2}}.$$
This formula is valid for a complete range of values of $\arg\lambda$ provided that $R(\tau)$ is negative when $t=t_2$; *i.e.* provided that the potential of $t_1$ is higher than the potential of $t_2$. Consequently the formula is valid for a complete or for an incomplete range of values of $\arg\lambda$ according as $x$ is outside or inside the continuous curve of Fig. 10.

*We shall finally shew that a branch of the line of force through $t_2$ either starts from* 1, *encircles $t_3$, and returns to* 1; *or else it starts from* 1 *and ends at* 0. The former is the case when $x$ is inside the dotted curve of Fig. 10, and the latter when $x$ is outside it.

First suppose that $x$ is inside the continuous curve of Fig. 10; then the equipotentials have the configuration of case (II). Consider the branch of the line of force which enters the horns of the closed crescent at $t_2$; it cannot cross the boundaries of the crescent without passing through an equilibrium point, and no such point exists; hence both ends of the branch must terminate at the point 1. Now the configuration of the lines of force only alters when $x$ crosses the dotted curve of Fig. 10. Hence, whenever $x$ is inside the dotted curve of Fig. 10, there is a branch of a line of force which starts from 1, goes to $t_2$, and returns to 1, obviously encircling $t_3$, which is in the region surrounded by the crescent.

When $x$ is outside the continuous curve of Fig. 10, the closed crescent contains the point 0 as well as 1; and the only possible change of configuration of the line of force is th[illegible] one of its ends† should be at 0 instead of both being at 1.

and of Clas

* It can only go near $t_3$ by assuming a form in which it passes through $t_2$, and we have just seen that it cannot assume this form when $x$ is inside the continuous curve of Fig. 10.

† Both ends cannot be at 0; for suppose the line charges replaced by surface distributions on circular cylinders of very small radius; since, by Gauss' theorem, the total charge inside the (closed) branch of the line of force under consideration is zero, and since the cha[illegible] curve, it would have also to contain at 0 (these charges being numeric[illegible] in sign). Therefore the quartic [illegible] forms part would have a cusp at [illegible] it from 0, it would bifurcate befo[illegible] neither of these events actually

merican Journal of Mathe- 221–249.

the Carnegie Institution of

To see that the branch of the line of force has one end at 0 and the other at 1, when $x$ is outside the dotted curve of Fig. 10, take $x$ large (greater than 9/8 is sufficient) and consider the limiting case when $x$ is positive, so that $\arg z$ is $\pm\pi$. In this case the nodal equipotentials coincide and form a curve consisting of two ovals crossing one another at $t_1$, $t_2$ (which are conjugate complexes); the left-hand oval contains the charges at 0 and $1/x$, and the right-hand oval contains the charges at $1/x$ and 1. And obviously the branch of the line of force through $t_2$ has its ends at 0 and 1; hence, whenever $R(x-9/8)$ is positive and $I(x)$ is very small, the line of force must pass very near 0; and so it must actually have its end at 0, in view of the manner in which lines of force radiate from the charge at 0.

Hence, whenever $x$ is outside the dotted curve of Fig. 10, a branch of a line of force passes from 0 to 1 by way of $t_2$; moreover $t_3$ lies in the region between this curve and the real axis; for $t_3$ is in the region surrounded by the closed crescent, and is consequently inside the region bounded by the line of force and any curve joining 0 and 1 and lying wholly inside the crescent; and, since $t_2$ and $t_3$ are on the same side of the real axis, the curve just mentioned is reconcilable with the real axis so far as $t_3$ is concerned. Hence, when $x$ is outside the dotted curve of Fig. 10, we get

$$I_6 \sim 2^{2\alpha+\beta-1}(-\sqrt{z})^{-\frac{1}{2}}(1-\sqrt{z})^{\gamma-\alpha-\beta}(3-\sqrt{z})^{1-\gamma}\left\{\frac{8(1-\sqrt{z})}{(3-\sqrt{z})^3}\right\}^{\lambda}\left(\frac{\pi}{\lambda}\right)^{\frac{1}{2}},$$

and it is easy to shew that $-\sqrt{z}$, $1-\sqrt{z}$, $3-\sqrt{z}$ have to be taken to have their arguments numerically less than $\pi$.

If, however, $x$ is inside the dotted curve of Fig. 10, the function which possesses the asymptotic expansion of which the dominant term has just been written down is

$$\int_1^{(1/x+)} t^{\beta+\lambda-1}(1-t)^{\gamma-\beta+2\lambda-1}(1-xt)^{-\alpha-\lambda}\,dt,$$

where the contour is described counter-clockwise or clockwise according as $I(x) \gtrless 0$.

By writing $t = 1 - u(x-1)/x$, this is easily found to be

$$\pm 2\pi i\frac{\Gamma(\gamma-\beta+2\lambda)\,e^{\pm\pi i(\gamma-\alpha-\beta-\lambda-1)}}{\Gamma(\gamma-\alpha-\beta+\lambda+1)\,\Gamma(\alpha+\lambda)}(1-x)^{\gamma-\alpha-\beta+\lambda}x^{\beta-\gamma-2\lambda}$$
$$\times F\left(\gamma-\beta+2\lambda,\ 1-\beta-\lambda;\ \gamma-\alpha-\beta+\lambda+1;\ \frac{x-1}{x}\right),$$

and so we get

$$\frac{\pi\Gamma(\gamma-\beta+2\lambda)}{\Gamma(\gamma-\alpha-\beta+\lambda+1)\,\Gamma(\alpha+\lambda)}(1-x)^{\gamma-\alpha-\beta+\lambda}x^{\beta-\gamma-2\lambda}$$
$$\times F\left(\gamma-\beta+2\lambda,\ 1-\beta-\lambda;\ \gamma-\alpha-\beta+1;\ \frac{x-1}{x}\right)$$
$$\sim 2^{2\alpha+\beta-2}z^{-\frac{1}{4}}(\sqrt{z}-1)^{\gamma-\alpha-\beta}(3-\sqrt{z})^{1-\gamma}\left\{\frac{8(\sqrt{z}-1)}{(3-\sqrt{z})^3}\right\}^{\lambda}\left(\frac{\pi}{\lambda}\right)^{\frac{1}{2}},$$

where $z = 9-8x$. By considering the potentials of $t_1$ and $t_2$, we see that, when the first type of asymptotic expansion (viz. that involved in $I_5$) is valid for a complete range of values of $\arg\lambda$, the second type is valid for an incomplete range, and *vice versa*.

# XV. *Asymptotic Satellites near the Equilateral-Triangle Equilibrium Points in the Problem of Three Bodies.*

By Professor DANIEL BUCHANAN, Queen's University, Kingston, Canada.

[*Received* 30 March, 1918. *Presented by* Professor Baker.]

## 1. INTRODUCTION.

If two finite bodies are subject to the Newtonian law of attraction and move in circles about their common centre of gravity, then there are five points, as Lagrange has shown*, at which an infinitesimal body would remain fixed with respect to the moving system if it were given an initial projection so as to be instantaneously fixed with respect to the finite bodies. Three of these points are situated on the line joining the finite bodies and these are called the *straight line* equilibrium points of the problem of three bodies. The remaining two points are situated at the vertices of the equilateral triangles having the line joining the finite masses as base. These points are called the *equilateral-triangle* equilibrium points of the problem of three bodies.

If the infinitesimal body is given a slight displacement from one of these points of equilibrium, and initial conditions are so determined that it will move in an orbit which is closed relatively to the moving system, it is called an *oscillating satellite.* If the infinitesimal body is disturbed slightly from an equilibrium point or from the periodic orbit about the equilibrium point, and initial conditions are so chosen that it will approach the equilibrium point or the periodic orbit, respectively, as the time approaches infinity, it will be called an *asymptotic satellite.*

The orbits which are asymptotic to the straight line equilibrium points were determined by Warren† in 1913. Those which are asymptotic to the periodic oscillations about these equilibrium points have been determined by the author of the present paper and appear in another memoir. [*Proc. Lon. Math. Soc.*, vol. XVII. (1918), p. 54.]

The periodic orbits to which they are asymptotic are the orbits of Class A and of Class B as determined by Moulton in chapter V. of his *Periodic Orbits*‡.

* Lagrange, *Collected Works*, vol. VI. pp. 229–324. Tisserand, *Mécanique Céleste*, vol. I. chap. VIII. Moulton, *Introduction to Celestial Mechanics* (New Edition, chap. VIII.).

† Warren, "A Class of Asymptotic Orbits in the Problem of Three Bodies," *American Journal of Mathematics*, vol. XXXVIII, No. 3, pp. 221–249.

‡ Publication No. 161 of the Carnegie Institution of Washington.

The object of the present paper is to make the discussion for the equilateral-triangle points of equilibrium which corresponds to the two papers mentioned above. The periodic orbits of the oscillating satellite which are approached by the asymptotic satellite are those determined by Buck*.

Two classes of periodic orbits are determined in Buck's memoir. One class of orbits is of two dimensions and lies wholly in the plane of motion of the finite bodies. This class exists only when one of the finite bodies is relatively small in comparison with the other body. The other class of orbits is of three dimensions, but there is no restriction as to the relative masses of the finite bodies.

The treatment of the problem under consideration is divided into two parts, Part I being devoted to asymptotic orbits of two dimensions, and Part II to asymptotic orbits of three dimensions.

The orbits considered in Part I are asymptotic to the equilibrium points themselves and not to the two-dimensional periodic orbits about these points. These asymptotic orbits exist only when the masses are more nearly equal than in the case of the two-dimensional periodic orbits. It is therefore doubtful if orbits exist which are asymptotic to the two-dimensional periodic orbits.

The three-dimensional orbits considered in Part II are asymptotic to the three-dimensional periodic orbits. The same restriction as to the relative masses of the finite bodies must be applied here as in Part I.

Only the formal constructions of the asymptotic solutions are made in this memoir. It has been shown, however, by Poincaré†, that if certain divisors which appear in the construction of such solutions do not vanish, then the solutions will converge for all values of the time $t$. Now $t$ can occur explicitly in the solutions only when such divisors vanish and, further, if $t$ does not occur explicitly then these divisors are different from zero. Hence, if the solutions can be constructed so that $t$ does not occur explicitly their convergence is assured, by Poincaré's theorem.

## PART I.

### Two-Dimensional Asymptotic Orbits.

#### 2. The Differential Equations.

Let the motion of the infinitesimal body be referred to a set of rotating rectangular axes $\xi$, $\eta$, $\zeta$, of which the origin is at the centre of mass of the finite bodies, and the $\xi\eta$-plane is the plane of their motion. The $\xi$ and $\eta$ axes rotate in the same direction as the finite bodies and with the same angular velocity. The masses of the finite bodies will be denoted by $\mu$ and $1-\mu$,

* Buck, "Oscillating Satellites near the Lagrangian Equilateral-Triangle Points," Moulton's *Periodic Orbits*, chap. IX. This paper will be cited as "Oscillating Satellite."

† *Mécanique Céleste*, vol. I. p. 341.

and the notation chosen so that $\mu \leqq \frac{1}{2}$. The distance between the finite bodies will be chosen as the unit of length, and the unit of time will be so chosen that the gravitational constant is unity.

If the coordinates of the infinitesimal body are denoted by $\xi$, $\eta$, and $\zeta$, and if derivatives with respect to $t$ are denoted by primes, then the differential equations of motion for the infinitesimal body are*

$$\left.\begin{aligned} \xi'' - 2\eta' = \frac{\partial U}{\partial \xi}, \quad \eta'' + 2\xi' = \frac{\partial U}{\partial \eta}, \quad \zeta'' = \frac{\partial U}{\partial \zeta}, \\ U = \tfrac{1}{2}(\xi^2 + \eta^2) + \frac{1-\mu}{\rho_1} + \frac{\mu}{\rho_2}, \\ \rho_1 = \sqrt{(\xi+\mu)^2 + \eta^2 + \zeta^2}, \quad \rho_2 = \sqrt{(\xi - 1 + \mu)^2 + \eta^2 + \zeta^2}, \end{aligned}\right\} \quad \text{..................(1)}$$

$\rho_1$ and $\rho_2$ being the distances from the infinitesimal body to the bodies $1-\mu$ and $\mu$ respectively.

The points of equilibrium are the solutions of the equations†

$$\frac{\partial U}{\partial \xi} = \frac{\partial U}{\partial \eta} = \frac{\partial U}{\partial \zeta} = 0.$$

There are two sets of points which satisfy these equations, but those with which we are concerned in this paper are

$$\text{I.} \quad \xi_0 = \tfrac{1}{2} - \mu, \quad \eta_0 = +\tfrac{1}{2}\sqrt{3}, \quad \zeta_0 = 0,$$

$$\text{II.} \quad \xi_0 = \tfrac{1}{2} - \mu, \quad \eta_0 = -\tfrac{1}{2}\sqrt{3}, \quad \zeta_0 = 0.$$

These two points lie in the rotating plane and at the vertices of the equilateral triangles having the line joining $1-\mu$ and $\mu$ as base. Obviously, the coordinates of the points differ only in the sign of $\sqrt{3}$. The asymptotic orbits will be discussed only for the point I, for on changing the sign of $\sqrt{3}$ we may obtain the corresponding results for the point II.

Let the origin be transferred to the point I by the transformation

$$\xi = \tfrac{1}{2} - \mu + \bar{\bar{x}}, \quad \eta = +\tfrac{1}{2}\sqrt{3} + \bar{\bar{y}}, \quad \zeta = \bar{\bar{z}}. \quad \text{..............................(2)}$$

Then the right members of the differential equations (1) can be expanded as power series in the new variables $\bar{\bar{x}}$, $\bar{\bar{y}}$, and $\bar{\bar{z}}$. These expansions converge only up to the singularities of the functions $1/\rho_1$ and $1/\rho_2$, that is, in the region which is common to the two spheres having their centres at the finite bodies and radii $\sqrt{2}$, but which excludes their centres.

Let a parameter $\epsilon$ be introduced into the differential equations by the substitutions

$$\bar{\bar{x}} = \epsilon x, \quad \bar{\bar{y}} = \epsilon y, \quad \bar{\bar{z}} = \epsilon z, \quad \text{........................................(3)}$$

where $x$, $y$, and $z$ are the new dependent variables. Then as a consequence of (2) and (3) the differential equations (1) become‡

$$\left.\begin{aligned} x'' - 2y' &= X_1 + X_2\epsilon + X_3\epsilon^2 + \dots, \\ y'' + 2x' &= Y_1 + Y_2\epsilon + Y_3\epsilon^2 + \dots, \\ z'' &= Z_1 + Z_2\epsilon + Z_3\epsilon^2 + \dots, \end{aligned}\right\} \quad \text{..............................(4)}$$

* Moulton, *Celestial Mechanics*, p. 280.

† *Ibid.* p. 290. Charlier, *Die Mechanik des Himmels*, vol. XI. pp. 102–111.

‡ "Oscillating Satellite," equations (4).

where

$$\left.\begin{aligned}
X_1 &= \tfrac{3}{4}[x+\sqrt{3}(1-2\mu)y], \quad Y_1 = \tfrac{3}{4}[\sqrt{3}(1-2\mu)x + \tfrac{3}{4}y],\\
X_2 &= +\tfrac{3}{16}[7(1-2\mu)x^2 + 2\sqrt{3}xy - 11(1-2\mu)y^2 + 4(1-2\mu)z^2],\\
Y_2 &= -\tfrac{3}{16}[\sqrt{3}x^2 + 22(1-2\mu)xy + 3\sqrt{3}y^2 - 4\sqrt{3}z^2],\\
Z_1 &= -z, \quad Z_2 = \tfrac{3}{2}[(1-2\mu)xz + \sqrt{3}yz],\\
X_3 &= \tfrac{1}{32}[-37x^3 + 75\sqrt{3}(1-2\mu)x^2y + 123xy^2 + 45\sqrt{3}(1-2\mu)y^3 - 12xz^2 + 6\sqrt{3}(1-2\mu)yz^2],\\
Y_3 &= \tfrac{1}{32}[-25\sqrt{3}(1-2\mu)x^3 + 123x^2y + 135\sqrt{3}(1-2\mu)xy^2 + 3y^3 - 60\sqrt{3}(1-2\mu)xz^2 + 132yz^2],\\
Z_3 &= -\tfrac{3}{8}[x^2z + 11y^2z - 4z^3 + 10\sqrt{3}(1-2\mu)xyz].
\end{aligned}\right\} \quad (5)$$

The remaining $X_n$, $Y_n$, and $Z_n$ are polynomials of degree $n$ in $x$, $y$, and $z$.

### 3. THE CHARACTERISTIC EXPONENTS.

For $\epsilon = 0$ equations (4) become

$$\left.\begin{aligned}
x'' - 2y' - \tfrac{3}{4}x - \tfrac{3}{4}\sqrt{3}(1-2\mu)y &= 0,\\
y'' + 2x' - \tfrac{3}{4}\sqrt{3}(1-2\mu)x - \tfrac{9}{4}y &= 0,\\
z'' + z &= 0.
\end{aligned}\right\} \quad \text{.............................. (6)}$$

The first two equations of (6) are independent of the last equation and can be integrated by putting

$$x = Ke^{\lambda t}, \quad y = Le^{\lambda t},$$

where $K$ and $L$ are arbitrary constants. The characteristic equation for the determination of $\lambda$ is

$$\lambda^4 + \lambda^2 + \tfrac{27}{4}\mu(1-\mu) = 0,$$

and the resulting values of $\lambda^2$ are

$$\lambda^2 = \frac{-1 \pm \sqrt{1 - 27\mu(1-\mu)}}{2}. \quad \text{.............................. (7)}$$

For small values of $\mu$ the expression under the radical is positive and numerically less than unity, and therefore the four values of $\lambda$ are purely imaginary. The limiting value of $\mu$ for which $\lambda$ is purely imaginary is that solution of

$$1 - 27\mu(1-\mu) = 0$$

which does not exceed $\frac{1}{2}$. This value is found to be

$$\mu = \mu_0 = \cdot 0385 \ldots.$$

For $\mu > \mu_0$ the four values of $\lambda$ are complex.

In order to construct asymptotic solutions of (4) it is necessary that at least one characteristic exponent shall be real or complex. Now in problems of dynamics in which the differential equations of motion do not involve $t$ explicitly, the characteristic exponents occur in pairs which differ only in sign*. Hence, if one value of $\lambda$ is real or complex, there must be another real or complex characteristic exponent which differs from the former only in sign. Further, if one pair of exponents is real or complex then $\mu \geqq \mu_0$, and it follows from (7) that the other pair of characteristic exponents is also real or complex, respectively. In view of the fact that the two-

* Poincaré, *Mécanique Céleste*, vol. I. chap. IV. p. 69.

dimensional periodic orbits exist only when $\mu < \mu_0$, it is impossible, according to the methods developed in this paper, to construct orbits which are asymptotic to the two-dimensional periodic orbits. We consider in Part I, therefore, the orbits which are asymptotic to the equilibrium points I and II and not to the two-dimensional periodic orbits.

For $\mu > \mu_0$ the characteristic exponents are $\sigma$, $-\sigma$, $\bar{\sigma}$, and $-\bar{\sigma}$, where

$$\left.\begin{aligned} \sigma &= \alpha + i\beta, \quad \bar{\sigma} = \alpha - i\beta, \\ \alpha &= \tfrac{1}{2}[\sqrt{27\mu(1-\mu)} - 1]^{\frac{1}{2}}, \\ \beta &= \tfrac{1}{2}[\sqrt{27\mu(1-\mu)} + 1]^{\frac{1}{2}}. \end{aligned}\right\} \qquad (8)$$

The quantities $\sigma$ and $\bar{\sigma}$, like most of the constants in the sequel, are conjugate complex. The notation adopted is such that a symbol having a stroke over it is the conjugate of the same symbol without the stroke.

## 4. Construction of Asymptotic Solutions.

We shall construct in (A) the solutions of (4) which approach zero as $t$ approaches $+\infty$. In (B) we shall show that the corresponding solutions which approach zero as $t$ approaches $-\infty$ may be obtained directly from those obtained in (A).

According to Poincaré's* definition of an asymptotic solution, each term must have the form

$$e^{\lambda t} P(t),$$

where $\lambda$ is a constant and $P$ is a periodic function of $t$. The solutions approach zero as $t$ approaches $-\infty$ or $+\infty$ according as the real part of $\lambda$ is positive or negative, respectively.

### (A) *Solutions in* $e^{-\alpha t}$.

In making the construction of asymptotic solutions it is convenient, although not necessary, to transform into the normal form the terms of the first two equations of (4) which are independent of $\epsilon$. In order to obtain the transformation for the introduction of normal variables, it is necessary to know the solutions of the first two equations of (6). They are

$$\left.\begin{aligned} x &= a_1 e^{\sigma t} + a_2 e^{-\sigma t} + a_3 e^{\bar{\sigma} t} + a_4 e^{-\bar{\sigma} t}, \\ y &= b_1 a_1 e^{\sigma t} + b_2 a_2 e^{-\sigma t} + \bar{b}_1 a_3 e^{\bar{\sigma} t} + \bar{b}_2 a_4 e^{-\bar{\sigma} t}, \end{aligned}\right\} \qquad (9)$$

where $a_1, \ldots, a_4$ are the constants of integration, and $b_1, \ldots, \bar{b}_2$ are constants so determined that (9) shall satisfy the first two equations of (6). It is found that

$$\left.\begin{aligned} b_1 &= \frac{8\sigma - 3\sqrt{3}(1-2\mu)}{9 - 4\sigma^2}, \\ b_2 &= \frac{8\sigma + 3\sqrt{3}(1-2\mu)}{4\sigma^2 - 9}. \end{aligned}\right\} \qquad (10)$$

Normal variables $x_1$, $x_2$, $x_3$, and $x_4$ are introduced by the substitutions†

$$\left.\begin{aligned} x &= x_1 + x_2 + x_3 + x_4, \\ y &= b_1 x_1 + b_2 x_2 + \bar{b}_1 x_3 + \bar{b}_2 x_4, \\ x' &= \sigma(x_1 - x_2) + \bar{\sigma}(x_3 - x_4), \\ y' &= \sigma(b_1 x_1 - b_2 x_2) + \bar{\sigma}(\bar{b}_1 x_3 - \bar{b}_2 x_4). \end{aligned}\right\} \qquad (11)$$

* Poincaré, *Mécanique Céleste*, vol. I. p. 340.

† *Ibid.* vol. I. p. 336.

The differential equations (4) then become

$$\left.\begin{aligned} x_1' - \sigma x_1 &= A_1(X_2\epsilon + X_3\epsilon^2 + \ldots) + B_1(Y_2\epsilon + Y_3\epsilon^2 + \ldots),\\ x_2' + \sigma x_2 &= A_2(X_2\epsilon + X_3\epsilon^2 + \ldots) + B_2(Y_2\epsilon + Y_3\epsilon^2 + \ldots),\\ x_3' - \bar{\sigma} x_3 &= A_3(X_2\epsilon + X_3\epsilon^2 + \ldots) + B_3(Y_2\epsilon + Y_3\epsilon^2 + \ldots),\\ x_4' + \bar{\sigma} x_4 &= A_4(X_2\epsilon + X_3\epsilon^2 + \ldots) + B_4(Y_2\epsilon + Y_3\epsilon^2 + \ldots), \end{aligned}\right\} \quad \ldots\ldots\ldots\ldots(12)$$

where

$$A_k = \frac{\Delta_{3k}}{\Delta}, \quad B_k = \frac{\Delta_{4k}}{\Delta}, \quad (k = 1, 2, 3, 4), \quad \ldots\ldots\ldots\ldots(13)$$

$\Delta$ being the determinant of the transformation (11), and $\Delta_{jk}$ the minors of the elements of this determinant, $j$ referring to the row and $k$ to the column. The computation shows that

$$A_3 = \bar{A}_1, \quad A_4 = \bar{A}_2, \quad B_3 = \bar{B}_1, \quad B_4 = \bar{B}_2.$$

The equations (12) will now be integrated as power series in $\epsilon$ by the method of undetermined coefficients. Accordingly let

$$\left.\begin{aligned} x_k &= x_k^{(0)} + x_k^{(1)}\epsilon + \ldots + x_k^{(j)}\epsilon^j + \ldots, \quad (k = 1, 2, 3, 4),\\ z &= z^{(0)} + z^{(1)}\epsilon + \ldots + z^{(j)}\epsilon^j + \ldots. \end{aligned}\right\} \quad \ldots\ldots\ldots\ldots(14)$$

Then from (11) it follows that $x$ and $y$ are likewise power series in $\epsilon$ of the form

$$x = \sum_{j=0}^{\infty} x^{(j)}\epsilon^j, \quad y = \sum_{j=0}^{\infty} y^{(j)}\epsilon^j, \quad \ldots\ldots\ldots\ldots(15)$$

where

$$\left.\begin{aligned} x^{(j)} &= x_1^{(j)} + x_2^{(j)} + x_3^{(j)} + x_4^{(j)},\\ y^{(j)} &= b_1 x_1^{(j)} + b_2 x_2^{(j)} + \bar{b}_1 x_3^{(j)} + \bar{b}_2 x_4^{(j)}. \end{aligned}\right\} \quad \ldots\ldots\ldots\ldots(16)$$

When equations (14) are substituted in (12) and the coefficients of the same powers of $\epsilon$ are equated in the resulting equations, we obtain sets of differential equations which define the various coefficients of $\epsilon$ in (14). In order that the solutions of these equations shall be asymptotic we impose the condition ($C_1$) that each term shall contain the factor $e^{-k\sigma t}$ or $e^{-k\bar{\sigma}t}$, where $k$ is a positive integer. This condition disposes of the two constants of integration which are associated with the exponentials $e^{\sigma t}$ and $e^{\bar{\sigma}t}$. There still remain the two constants associated with the other exponentials $e^{-\sigma t}$ and $e^{-\bar{\sigma}t}$, and in order that these shall be uniquely determined we impose the conditions ($C_2$) that

$$x = a, \quad y = 0,$$

at $t = 0$, that is, we suppose that the infinitesimal body is initially displaced from the point I at the distance $a$ on the $x$-axis. When these conditions are imposed on (15) we obtain

$$x^{(0)}(0) = a, \quad x^{(j)}(0) = 0, \quad (j = 1, \ldots \infty),$$
$$y^{(j)}(0) = 0, \quad (j = 0, \ldots \infty).$$

The terms which are independent of $\epsilon$ when equations (14) are substituted in (12) are

$$\left.\begin{aligned} x_1'^{(0)} - \sigma x_1^{(0)} = 0, \quad & x_2'^{(0)} + \sigma x_2^{(0)} = 0,\\ x_3'^{(0)} - \bar{\sigma} x_3^{(0)} = 0, \quad & x_4'^{(0)} + \bar{\sigma} x_4^{(0)} = 0,\\ z''^{(0)} + z^{(0)} &= 0. \end{aligned}\right\} \quad \ldots\ldots\ldots\ldots(17)$$

The solutions of these equations are

$$\left.\begin{aligned} x_1^{(0)} = d_1^{(0)} e^{\sigma t}, \quad x_2^{(0)} = d_2^{(0)} e^{-\sigma t}, \quad x_3^{(0)} = d_3^{(0)} e^{\bar{\sigma} t}, \quad x_4^{(0)} = d_4^{(0)} e^{-\bar{\sigma} t},\\ z^{(0)} = d_5^{(0)} \sin t + d_6^{(0)} \cos t, \end{aligned}\right\} \quad \ldots\ldots\ldots\ldots(18)$$

where $d_1^{(0)}, \ldots, d_6^{(0)}$ are the constants of integration. From condition ($C_1$) it follows that all these constants are zero except $d_2^{(0)}$ and $d_4^{(0)}$. When the resulting values of (18) are substituted in (16) we obtain

$$\left.\begin{aligned} x^{(0)} &= d_2^{(0)} e^{-\sigma t} + d_4^{(0)} e^{-\bar{\sigma} t}, \\ y^{(0)} &= b_2 d_2^{(0)} e^{-\sigma t} + \bar{b}_2 d_4^{(0)} e^{-\bar{\sigma} t}. \end{aligned}\right\} \quad \ldots\ldots\ldots\ldots\ldots\ldots(19)$$

From conditions ($C_2$) it follows that

$$d_2^{(0)} + d_4^{(0)} = a, \quad b_2 d_2^{(0)} + \bar{b}_2 d_4^{(0)} = 0,$$

then

$$d_2^{(0)} = a d^{(0)}, \quad d_4^{(0)} = \bar{d}_2^{(0)} = a \bar{d}^{(0)},$$

where

$$d^{(0)} = \frac{\bar{b}_2}{\bar{b}_2 - b_2}.$$

With this determination of the arbitrary constants $d_2^{(0)}$ and $d_4^{(0)}$, the solutions (19) take the form

$$\left.\begin{aligned} x^{(0)} &= a\left[d^{(0)} e^{-\sigma t} + \bar{d}^{(0)} e^{-\bar{\sigma} t}\right], \\ y^{(0)} &= a\left[b_2 d^{(0)} e^{-\sigma t} + \bar{b}_2 \bar{d}^{(0)} e^{-\bar{\sigma} t}\right]. \end{aligned}\right\} \quad \ldots\ldots\ldots\ldots\ldots\ldots(20)$$

The differential equations obtained by equating the coefficients of $\epsilon$ in (12), after equations (14) are substituted, are

$$\left.\begin{aligned} x_1'^{(1)} - \sigma x_1^{(1)} = X_1^{(1)}, &\quad x_2'^{(1)} + \sigma x_2^{(1)} = X_2^{(1)}, \\ x_3'^{(1)} - \bar{\sigma} x_3^{(1)} = X_3^{(1)}, &\quad x_4'^{(1)} + \bar{\sigma} x_4^{(1)} = X_4^{(1)}, \\ z''^{(1)} + z^{(1)} &= 0, \end{aligned}\right\} \quad \ldots\ldots\ldots\ldots\ldots\ldots(21)$$

where

$$X_1^{(1)} = a^2\left[M_{20}^{(1)} e^{-2\sigma t} + M_{11}^{(1)} e^{-(\sigma+\bar{\sigma})t} + M_{02}^{(1)} e^{-2\bar{\sigma} t}\right],$$
$$X_2^{(1)} = a^2\left[N_{20}^{(1)} e^{-2\sigma t} + N_{11}^{(1)} e^{-(\sigma+\bar{\sigma})t} + N_{02}^{(1)} e^{-2\bar{\sigma} t}\right],$$
$$X_3^{(1)} = \bar{X}_1^{(1)}, \quad X_4^{(1)} = \bar{X}_2^{(1)},$$
$$M_{20}^{(1)} = A_1 M_0 + B_1 N_0, \quad N_{20}^{(1)} = A_2 M_0 + B_2 N_0,$$
$$M_{02}^{(1)} = A_1 \bar{M}_0 + B_1 \bar{N}_0, \quad N_{02}^{(1)} = A_2 \bar{M}_0 + B_2 \bar{N}_0,$$
$$M_{11}^{(1)} = A_1 M_1 + B_1 N_1, \quad N_{11}^{(1)} = A_2 M_1 + B_2 N_1,$$
$$M_0 = +\tfrac{3}{16} a^2 (d^{(0)})^2 \left[7(1-2\mu) + 2\sqrt{3} b_2 - 11(1-2\mu) b_2^2\right],$$
$$N_0 = -\tfrac{3}{16} a^2 (d^{(0)})^2 \left[\sqrt{3} + 22(1-2\mu) b_2 + 3\sqrt{3} b_2^2\right],$$
$$M_1 = +\tfrac{3}{8} a^2 d^{(0)} \bar{d}^{(0)} \left[7(1-2\mu) + \sqrt{3}(b_2 + \bar{b}_2) - 11(1-2\mu) b_2 \bar{b}_2\right],$$
$$N_1 = -\tfrac{3}{8} a^2 d^{(0)} \bar{d}^{(0)} \left[\sqrt{3} + 11(1-2\mu)(b_2 + \bar{b}_2) + 3\sqrt{3} b_2 \bar{b}_2\right].$$

The constants $A_1$, $A_2$, $B_1$, and $B_2$ have the values defined in equations (13).

The general solutions of (21) are

$$\left.\begin{aligned} x_1^{(1)} &= d_1^{(1)} e^{\sigma t} + a^2\left[m_{20}^{(1)} e^{-2\sigma t} + m_{11}^{(1)} e^{-(\sigma+\bar{\sigma})t} + m_{02}^{(1)} e^{-2\bar{\sigma} t}\right], \\ x_2^{(1)} &= d_2^{(1)} e^{-\sigma t} + a^2\left[n_{20}^{(1)} e^{-2\sigma t} + n_{11}^{(1)} e^{-(\sigma+\bar{\sigma})t} + n_{02}^{(1)} e^{-2\bar{\sigma} t}\right], \\ x_3^{(1)} &= d_3^{(1)} e^{\bar{\sigma} t} + a^2\left[\bar{m}_{02}^{(1)} e^{-2\sigma t} + \bar{m}_{11}^{(1)} e^{-(\sigma+\bar{\sigma})t} + \bar{m}_{20}^{(1)} e^{-2\bar{\sigma} t}\right], \\ x_4^{(1)} &= d_4^{(1)} e^{-\bar{\sigma} t} + a^2\left[\bar{n}_{02}^{(1)} e^{-2\sigma t} + \bar{n}_{11}^{(1)} e^{-(\sigma+\bar{\sigma})t} + \bar{n}_{20}^{(1)} e^{-2\bar{\sigma} t}\right], \\ z^{(1)} &= d_5^{(1)} \sin t + d_6^{(1)} \cos t, \end{aligned}\right\} \quad \ldots\ldots\ldots\ldots(22)$$

where $d_1^{(1)}, \ldots, d_6^{(1)}$ are the constants of integration and

$$\left.\begin{aligned} m_{20}^{(1)} &= -\frac{M_{20}^{(1)}}{3\sigma}, \quad m_{11}^{(1)} = -\frac{M_{11}^{(1)}}{2\sigma+\bar\sigma}, \quad m_{02}^{(1)} = -\frac{M_{02}^{(1)}}{2\bar\sigma+\sigma}, \\ n_{20}^{(1)} &= -\frac{N_{20}^{(1)}}{\sigma}, \quad n_{11}^{(1)} = -\frac{N_{11}^{(1)}}{\bar\sigma}, \quad n_{02}^{(1)} = -\frac{N_{02}^{(1)}}{2\bar\sigma-\sigma}. \end{aligned}\right\} \quad \ldots\ldots\ldots\ldots\ldots(23)$$

From condition ($C_1$) it follows that $d_1^{(1)} = d_3^{(1)} = d_5^{(1)} = d_6^{(1)} = 0$. Then, on substituting (22) in (16), we obtain

$$\left.\begin{aligned} x^{(1)} &= d_2^{(1)} e^{-\sigma t} + d_4^{(1)} e^{-\bar\sigma t} + a^2 [A_{20}^{(1)} e^{-2\sigma t} + A_{11}^{(1)} e^{-(\sigma+\bar\sigma)t} + A_{02}^{(1)} e^{-2\bar\sigma t}], \\ y^{(1)} &= b_2 d_2^{(1)} e^{-\sigma t} + \bar b_2 d_4^{(1)} e^{-\bar\sigma t} + a^2 [B_{20}^{(1)} e^{-2\sigma t} + B_{11}^{(1)} e^{-(\sigma+\bar\sigma)t} + B_{02}^{(1)} e^{-2\bar\sigma t}], \\ z^{(1)} &= 0, \end{aligned}\right\} \quad \ldots\ldots(24)$$

where

$$\begin{aligned} A_{20}^{(1)} &= m_{20}^{(1)} + n_{20}^{(1)} + \bar m_{02}^{(1)} + \bar n_{02}^{(1)}, \\ A_{11}^{(1)} &= m_{11}^{(1)} + n_{11}^{(1)} + \bar m_{11}^{(1)} + \bar n_{11}^{(1)}, \\ B_{20}^{(1)} &= b_1 m_{20}^{(1)} + b_2 n_{20}^{(1)} + \bar b_1 \bar m_{02}^{(1)} + \bar b_2 \bar n_{02}^{(1)}, \\ B_{11}^{(1)} &= b_1 m_{11}^{(1)} + b_2 n_{11}^{(1)} + \bar b_1 \bar m_{11}^{(1)} + \bar b_2 \bar n_{11}^{(1)}, \\ A_{02}^{(1)} &= \bar A_{20}^{(1)}, \quad B_{02}^{(1)} = \bar B_{20}^{(1)}, \quad A_{11}^{(1)} = \bar A_{11}^{(1)}, \quad B_{11}^{(1)} = \bar B_{11}^{(1)}. \end{aligned}$$

It is observed that the constants $A_{11}^{(1)}$ and $B_{11}^{(1)}$ are real. The only undetermined constants in (24) are $d_2^{(1)}$ and $d_4^{(1)}$. When conditions ($C_2$) are imposed, it is found that these constants can be uniquely determined and that they are conjugates. Since they carry the factor $a^2$ let

$$d_2^{(1)} = a^2 A_{10}^{(1)}, \quad d_4^{(1)} = a^2 A_{01}^{(1)}, \quad b_2 d_2^{(1)} = a^2 B_{10}^{(1)}, \quad \bar b_2 d_4^{(1)} = a^2 B_{01}^{(1)}:$$

then the solutions (24) may be expressed in the form

$$\left.\begin{aligned} x^{(1)} &= a^2 \sum_{j=0}^{2} \sum_{k=0}^{2} A_{jk}^{(1)} e^{-(j\sigma+k\bar\sigma)t}, \quad 0 < j+k \leqq 2, \\ y^{(1)} &= a^2 \sum_{j=0}^{2} \sum_{k=0}^{2} B_{jk}^{(1)} e^{-(j\sigma+k\bar\sigma)t}, \\ z^{(1)} &= 0. \end{aligned}\right\} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots(25)$$

The coefficients have the property that

$$A_{jk}^{(1)} = \bar A_{kj}^{(1)}, \quad B_{jk}^{(1)} = \bar B_{kj}^{(1)}.$$

If $j \neq k$ these coefficients are conjugate complex, but if $j = k$ they are real.

The remaining steps of the integration can be carried on in an entirely similar way. In order to find the general term we proceed by induction.

Let us suppose that $x^{(h)}$, $y^{(h)}$, and $z^{(h)}$ have been determined for $h = 0, 1, \ldots, \nu-1$; and that

$$\left.\begin{aligned} x^{(h)} &= a^{h+1} \sum_{j=0}^{h+1} \sum_{k=0}^{h+1} A_{jk}^{(h)} e^{-(j\sigma+k\bar\sigma)t}, \quad 0 < j+k \leq h+1, \\ y^{(h)} &= a^{h+1} \sum_{j=0}^{h+1} \sum_{k=0}^{h+1} B_{jk}^{(h)} e^{-(j\sigma+k\bar\sigma)t}, \\ z^{(h)} &= 0, \end{aligned}\right\} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots(26)$$

where $A_{jk}^{(h)}$ and $B_{jk}^{(h)}$ are constants such that

$$A_{jk}^{(h)} = \bar A_{kj}^{(h)}, \quad B_{jk}^{(h)} = \bar B_{kj}^{(h)}.$$

If $j \neq k$ these constants are conjugate complex, but if $j = k$ they are real. We proceed to show that the solutions have the same properties when $h = \nu$.

The differential equations which define $x_1^{(\nu)}$, $x_2^{(\nu)}$, $x_3^{(\nu)}$, and $x_4^{(\nu)}$ are the same as (21) except the right members, while the equation in $z^{(\nu)}$ is the same as in (21). Let the right members be denoted by $X_1^{(\nu)}$, $X_2^{(\nu)}$, $X_3^{(\nu)}$, and $X_4^{(\nu)}$ respectively. These right members are functions of $x^{(1)}, \ldots, x^{(\nu-1)}$, which are assumed to be known. They have the form

$$\left.\begin{aligned} X_1^{(\nu)} &= a^{\nu+1}\left[\sum_{j=0}^{\nu-1}\sum_{k=0}^{\nu-1} M_{jk}^{(\nu)} e^{-(j\sigma+k\bar{\sigma})t}\right], \\ X_2^{(\nu)} &= a^{\nu+1}\left[\sum_{j=0}^{\nu+1}\sum_{k=0}^{\nu+1} N_{jk}^{(\nu)} e^{-(j\sigma+k\bar{\sigma})t}\right], \\ X_3^{(\nu)} &= \bar{X}_1^{(\nu)}, \quad X_4^{(\nu)} = \bar{X}_2^{(\nu)}, \\ & 2 \leqq j+k \leqq \nu+1, \end{aligned}\right\} \qquad (27)$$

where the coefficients of the exponentials are known constants. Since $j+k \geqq 2$, the right members will contain no terms in $e^{-\sigma t}$ or $e^{-\bar{\sigma}t}$, and therefore the particular integrals will have the same form as the right members. The complementary functions are the same as at the previous steps, and therefore the complete solutions have the form

$$\left.\begin{aligned} x_1^{(\nu)} &= d_1^{(\nu)} e^{\sigma t} + a^{\nu+1}\sum_{j=0}^{\nu+1}\sum_{k=0}^{\nu+1} m_{jk}^{(\nu)} e^{-(j\sigma+k\bar{\sigma})t}, \\ x_2^{(\nu)} &= d_2^{(\nu)} e^{-\sigma t} + a^{\nu+1}\sum_{j=0}^{\nu+1}\sum_{k=0}^{\nu+1} n_{jk}^{(\nu)} e^{-(j\sigma+k\bar{\sigma})t}, \\ x_3^{(\nu)} &= d_3^{(\nu)} e^{\bar{\sigma} t} + a^{\nu+1}\sum_{j=0}^{\nu+1}\sum_{k=0}^{\nu+1} \bar{m}_{kj}^{(\nu)} e^{-(j\sigma+k\bar{\sigma})t}, \\ x_4^{(\nu)} &= d_4^{(\nu)} e^{-\bar{\sigma} t} + a^{\nu+1}\sum_{j=0}^{\nu+1}\sum_{k=0}^{\nu+1} \bar{n}_{kj}^{(\nu)} e^{-(j\sigma+k\bar{\sigma})t}, \\ z^{(\nu)} &= d_5^{(\nu)} \sin t + d_6^{(\nu)} \cos t, \end{aligned}\right\} \qquad (28)$$

where $d_1^{(\nu)}, \ldots, d_6^{(\nu)}$ are the constants of integration, and the remaining coefficients are known constants. From condition ($C_1$) it follows that

$$d_1^{(\nu)} = d_3^{(\nu)} = d_5^{(\nu)} = d_6^{(\nu)} = 0.$$

On substituting the resulting values of (28) in (16) and imposing conditions ($C_2$) that

$$x^{(\nu)}(0) = y^{(\nu)}(0) = 0,$$

it is found that $d_2^{(\nu)}$ and $d_4^{(\nu)}$ are uniquely determined. As they contain the factor $a^{\nu+1}$ and are conjugate complex, we may put

$$d_2^{(\nu)} = a^{\nu+1} A_{10}^{(\nu)}, \quad d_4^{(\nu)} = a^{\nu+1} A_{01}^{(\nu)}, \quad b_2 d_2^{(\nu)} = a^{\nu+1} B_{10}^{(\nu)}, \quad \bar{b}_2 d_4^{(\nu)} = a^{\nu+1} B_{01}^{(\nu)},$$

in which case the solutions for $x^{(\nu)}$ and $y^{(\nu)}$ are of the same form as (26) if $h = \nu+1$ in (26). This completes the induction.

Thus the integration of the differential equations (12) can be carried on to any degree of accuracy desired. It remains to be shown that the solutions which have been determined are real for real values of $a$. This will now be discussed.

Consider a typical term

$$A_{j_1 j_2}^{(h)} e^{-(j_1\sigma + j_2\bar{\sigma})t}, \quad 0 < j_1 + j_2 \leqq h+1, \qquad (29)$$

of the solution for $x^{(h)}$ in (26). If $j_1 = j_2 = j$, then $A_{jj}^{(h)}$ is real, and, since $\sigma = \alpha + i\beta$, the term (29) becomes $A_{jj}^{(h)}e^{-2j\alpha t}$, which is real. If $j_1 \neq j_2$, then there is associated with (29) the term

$$A_{j_2 j_1}^{(h)} e^{-(j_2\sigma+j_1\bar{\sigma})t}, \quad 0 < j_1 + j_2 \leqq h+1, \qquad (30)$$

in which $A_{j_1 j_2}^{(h)}$ and $A_{j_2 j_1}^{(h)}$ are conjugate. If we put

$$A_{j_1 j_2}^{(h)} = a^{(h)} + ib^{(h)}, \quad A_{j_2 j_1}^{(h)} = a^{(h)} - ib^{(h)},$$

then the sum of the terms (29) and (30) becomes

$$2e^{-(j_1+j_2)\alpha t}\left[a^{(h)} \cos (j_1 - j_2)\beta t + b^{(h)} \sin (j_1 - j_2)\beta t\right],$$

which is real. As (29) and (30) are typical of the terms in the solutions for both $x^{(h)}$ and $y^{(h)}$, these solutions are real. They may be expressed in the form

$$x^{(h)} = a^{h+1} \sum_{j=1}^{h+1} \sum_{k=0}^{j} e^{-j\alpha t}\left[a_{jk}^{(h)} \cos k\beta t + b_{jk}^{(h)} \sin k\beta t\right],$$

$$y^{(h)} = a^{h+1} \sum_{j=1}^{h+1} \sum_{k=0}^{j} e^{-j\alpha t}\left[c_{jk}^{(h)} \cos k\beta t + d_{jk}^{(h)} \sin k\beta t\right],$$

where the coefficients are real constants.

On substituting the above solutions in (15), and the results in (3), it is found that the solutions for $\bar{\bar{x}}$, $\bar{\bar{y}}$, and $\bar{\bar{z}}$ carry factors in $a$ and $\epsilon$, but only as products and to the same degrees. We may therefore put $a = 1$, and when the resulting values for $\bar{\bar{x}}$, $\bar{\bar{y}}$, and $\bar{\bar{z}}$ are substituted in (2) we obtain

$$\left.\begin{aligned} \xi &= \tfrac{1}{2} - \mu + \sum_{h=1}^{\infty} \sum_{j=1}^{h} \sum_{k=0}^{j} e^{-j\alpha t}\left[a_{jk}^{(h)} \cos k\beta t + b_{jk}^{(h)} \sin k\beta t\right] \epsilon^h, \\ \eta &= +\tfrac{1}{2}\sqrt{3} + \sum_{h=1}^{\infty} \sum_{j=1}^{h} \sum_{k=0}^{j} e^{-j\alpha t}\left[c_{jk}^{(h)} \cos k\beta t + d_{jk}^{(h)} \sin k\beta t\right] \epsilon^h, \\ \zeta &= 0, \end{aligned}\right\} \qquad (31, \text{I})$$

which represents the orbit that approaches the point I as $t$ approaches $+\infty$. By changing the sign of $\sqrt{3}$ in (31, I) we obtain the orbit which approaches the point II as $t$ approaches $+\infty$. This latter orbit will be referred to as (31, II). Obviously, both orbits (31, I) and (31, II) are of two dimensions and lie in the plane of motion of the finite bodies.

(B) *Solutions in* $e^{\alpha t}$.

Let us next consider the orbits which approach the points I and II as the time approaches $-\infty$. These solutions could be constructed in a manner entirely similar to the preceding construction, but they may be obtained more directly, as we shall show, from the former solutions.

Consider the differential equations (4). Obviously, the orbits under consideration are of two dimensions as in the former case, and the variable $\zeta$ may be suppressed. Since $U$ in equation (1) is even in $\eta$, it follows that $\frac{\partial U}{\partial \xi}$ is even in $\eta$ and $\frac{\partial U}{\partial \eta}$ is odd in $\eta$. When the substitutions (2) and (3) are made in (1) the parity of $\eta$, or its equivalent substitutions, is unaltered in the right members of the differential equations. Hence the right member of the first equation of (4) is even in $\eta$ and $\sqrt{3}$, considered together, while the right member of the second equation of (4) is odd in $\eta$ and $\sqrt{3}$, considered together. Let the solutions of (4) be denoted by

$$x = f_1(t, +\sqrt{3}), \quad y = f_2(t, +\sqrt{3}). \qquad (32)$$

Further, let $$t=-\tau, \quad y=-p$$ be substituted in (4) and the sign of $\sqrt{3}$ changed. We obtain differential equations (4′) which are identically the same in $x$, $p$ and $\tau$ as (4) are in $x$, $y$ and $t$. For the corresponding changes in the initial conditions, which, in fact, are unaltered by the change in sign of $t$, $y$ and $\sqrt{3}$, we obtain as the solutions of (4′)

$$x=f_1(\tau, -\sqrt{3}), \quad p=f_2(\tau, -\sqrt{3}),$$

where $f_1$ and $f_2$ are the same functions as in (32), but of different arguments, as indicated. On restoring the former variables, we find that

$$x=f_1(-t, -\sqrt{3}), \quad y=-f_2(-t, -\sqrt{3})$$

are solutions of (4), and that the equations

$$\left.\begin{aligned}\xi &= \tfrac{1}{2}-\mu+f_1(-t, -\sqrt{3}),\\ \eta &= +\tfrac{1}{2}\sqrt{3}-f_2(-t, -\sqrt{3})\end{aligned}\right\} \quad \text{..............................(33, I)}$$

represent the orbit which approaches the equilibrium point I as $t$ approaches $-\infty$. But (33, I) is obtained from (31, II) by changing the signs of $t$ and $\eta$ in (31, II). Thus the solution which approaches the point I as $t$ approaches $-\infty$ can be obtained from the solution which approaches the point II as $t$ approaches $+\infty$ by changing the signs of $t$ and $\eta$ in the latter. Further, the equations

$$\left.\begin{aligned}\xi &= \tfrac{1}{2}-\mu+f_1(-t, +\sqrt{3}),\\ \eta &= -\tfrac{1}{2}\sqrt{3}-f_2(-t, +\sqrt{3})\end{aligned}\right\} \quad \text{............................(33, II)}$$

represent the orbit which approaches the point II as $t$ approaches $-\infty$, and they may be obtained by changing the signs of $t$ and $\eta$ in the solutions which approach the point I as $t$ approaches $+\infty$. Thus the orbit which approaches one equilibrium point as $t$ approaches $+\infty$ or $-\infty$ may be obtained from the orbit which approaches the other equilibrium point as $t$ approaches $-\infty$ or $+\infty$, respectively, by changing the signs of $t$ and $\eta$ in the solutions for the other orbit.

## 5. Geometrical Considerations.

The parameter $\epsilon$ remains arbitrary in the asymptotic solutions. From the way in which the initial conditions ($C_2$) were chosen, it follows that $\epsilon$ denotes the initial displacement of the infinitesimal body from the equilibrium point and parallel to the $\xi$-axis. As the solutions contain $\epsilon$ both to even and odd degrees, the shape of the orbit will vary not only with the numerical value of $\epsilon$ but also with the sign of $\epsilon$.

The direction in which an orbit approaches an equilibrium point is indeterminate and therefore independent of $\epsilon$. In order to prove this we need to show that

$$\operatorname*{limit}_{t=\pm\infty} \frac{d\bar{y}}{d\bar{x}}$$

is indeterminate. We shall consider only the orbit which approaches the point I as $t$ approaches $+\infty$, since the discussion for the other orbits is essentially the same.

When $t$ becomes very large, the most important terms of $\bar{\bar{x}}$ and $\bar{\bar{y}}$ are those in $e^{-\alpha t}$ to the first degree. These are

$$\bar{\bar{x}} = \sum_{h=1}^{\infty} e^{-\alpha t} [a_{11}^{(h)} \cos \beta t + b_{11}^{(h)} \sin \beta t] \epsilon^h,$$

$$\bar{\bar{y}} = \sum_{h=1}^{\infty} e^{-\alpha t} [c_{11}^{(h)} \cos \beta t + d_{11}^{(h)} \sin \beta t] \epsilon^h,$$

where the coefficients are real constants. Now

$$\operatorname*{limit}_{t=+\infty} \frac{d\bar{\bar{y}}}{d\bar{\bar{x}}} = \operatorname*{limit}_{t=+\infty} \frac{d\bar{\bar{y}}}{dt} \Big/ \frac{d\bar{\bar{x}}}{dt} = \operatorname*{limit}_{t=+\infty} \frac{\sum_{h=1}^{\infty} [(\beta d_{11}^{(h)} - \alpha c_{11}^{(h)}) \cos \beta t - (\alpha d_{11}^{(h)} + \beta c_{11}^{(h)}) \sin \beta t] \epsilon^h}{\sum_{h=1}^{\infty} [(\beta b_{11}^{(h)} - \alpha a_{11}^{(h)}) \cos \beta t - (\alpha b_{11}^{(h)} + \beta a_{11}^{(h)}) \sin \beta t] \epsilon^h},$$

and this limit is indeterminate.

For a given value of $\epsilon$ there are two orbits approaching each equilibrium point, according as $t$ approaches $+\infty$ or $-\infty$. Since the equations (31, I) and (33, II) differ only in the signs of $t$ and $\eta$, and similarly with (31, II) and (33, I), it follows that the orbit which approaches one equilibrium point as $t$ approaches $-\infty$ has the same shape as the orbit which approaches the other equilibrium point as $t$ approaches $+\infty$. The two orbits which approach one equilibrium point for a given value of $\epsilon$ are the reflection in the $\xi$-axis of the two orbits which approach the other equilibrium point for the same value of $\epsilon$. This would be expected from the symmetrical nature of the problem.

### 5 *a*. NUMERICAL EXAMPLE.

To illustrate the nature of these asymptotic orbits we have assigned to $\mu$ the particular value 0·1 and have constructed the solutions which approach the equilibrium point I as $t$ approaches $+\infty$. The results are

$$\left.\begin{aligned} x &= e^{-\alpha t} (\cos \beta t + 0{\cdot}724 \sin \beta t) \epsilon + [e^{-\alpha t} (0{\cdot}045 \cos \beta t + 1{\cdot}002 \sin \beta t) \\ &\qquad + e^{-2\alpha t} (0{\cdot}267 - 0{\cdot}312 \cos 2\beta t - 0{\cdot}392 \sin 2\beta t)] \epsilon^2 + \ldots, \\ y &= e^{-\alpha t} (-0{\cdot}105 \sin \beta t) \epsilon + [e^{-\alpha t} (0{\cdot}669 \cos \beta t - 0{\cdot}532 \sin \beta t) \\ &\qquad + e^{-2\alpha t} (-0{\cdot}679 + 0{\cdot}010 \cos 2\beta t + 0{\cdot}120 \sin 2\beta t)] \epsilon^2 + \ldots, \end{aligned}\right\} \ldots(34)$$

where $\alpha = 0{\cdot}374$ and $\beta = 0{\cdot}800$.

If the infinitesimal body is projected from the positive $x$-axis at the initial distance 0·1 from the point I, then $\epsilon = 0{\cdot}1$, and the solutions (34) together with their derivatives become

$$\left.\begin{aligned} x &= e^{-\alpha t} (0{\cdot}1004 \cos \beta t + 0{\cdot}0824 \sin \beta t) \\ &\qquad + e^{-2\alpha t} (0{\cdot}0027 - 0{\cdot}0031 \cos 2\beta t - 0{\cdot}0039 \sin 2\beta t) + \ldots, \\ y &= e^{-\alpha t} (0{\cdot}0067 \cos \beta t - 0{\cdot}1105 \sin \beta t) \\ &\qquad + e^{-2\alpha t} (-0{\cdot}0068 + 0{\cdot}0001 \cos 2\beta t + 0{\cdot}0012 \sin 2\beta t) + \ldots, \\ x' &= e^{-\alpha t} (0{\cdot}0284 \cos \beta t - 0{\cdot}1111 \sin \beta t) \\ &\qquad + e^{-2\alpha t} (-0{\cdot}0020 - 0{\cdot}0039 \cos 2\beta t + 0{\cdot}0078 \sin 2\beta t) + \ldots, \\ y' &= e^{-\alpha t} (0{\cdot}0909 \cos \beta t + 0{\cdot}0359 \sin \beta t) \\ &\qquad + e^{-2\alpha t} (0{\cdot}0051 + 0{\cdot}0018 \cos 2\beta t - 0{\cdot}0011 \sin 2\beta t) + \ldots. \end{aligned}\right\} \ldots(35)$$

As $t$ increases, the most important terms of these solutions are those which carry $e^{-at}$ as a factor. On considering only these terms and examining the value of

$$\frac{d^2y}{dx^2} = \frac{y''x' - x''y'}{(x')^3},$$

we find that $\frac{d^2y}{dx^2}$ cannot change sign by passing through zero for any real value of $t$. Hence the curve represented by the first two equations of (35) is always concave to the origin, *i.e.* to the point I.

Numerical values have been computed for equations (35) and they are to be found in Table I.

TABLE I.

$\mu = 0{\cdot}1$ $\epsilon = 0{\cdot}1$

| $t$ | $x$ | $y$ | $x'$ | $y'$ | $\frac{dy}{dx} = \frac{y'}{x'}$ |
|---|---|---|---|---|---|
| 0 | + ·1000 | 0 | + ·0225 | − ·0840 | − 3·7 |
| ·1 | + ·1017 | − ·0082 | + ·0144 | − ·0782 | − 5·4 |
| ·2 | + ·1091 | − ·0157 | + ·0068 | − ·0723 | − 10·6 |
| ·3 | + ·1032 | − ·0225 | − ·0004 | − ·0663 | + 16·6 |
| ·4 | + ·1028 | − ·0288 | − ·0071 | − ·0603 | + 8·5 |
| ·5 | + ·1018 | − ·0346 | − ·0135 | − ·0542 | + 4·0 |
| 1 | + ·0883 | − ·0542 | − ·0384 | − ·0232 | + 0·6 |
| 1·5 | + ·0655 | − ·0594 | − ·0513 | + ·0013 | − 0·03 |
| 2 | + ·0389 | − ·0539 | − ·0526 | + ·0190 | − 0·4 |
| 2·5 | + ·0142 | − ·0418 | − ·0451 | + ·0284 | − 6·3 |
| 3 | − ·0054 | − ·0267 | − ·0323 | + ·0303 | − 9·4 |
| 3·5 | − ·0179 | − ·0122 | − ·0179 | + ·0269 | − 1·5 |
| ·4 | − ·0236 | − ·0004 | − ·0049 | + ·0202 | − 4·0 |
| 4·5 | − ·0234 | + ·0077 | + ·0044 | + ·0124 | + 2·8 |
| 5 | − ·0193 | + ·0121 | + ·0102 | + ·0049 | + 0·5 |
| 6 | − ·0077 | + ·0107 | + ·0121 | − ·0046 | − 0·4 |
| 7 | + ·0020 | + ·0054 | + ·0067 | − ·0068 | − 1·0 |
| 8 | + ·0055 | − ·0003 | + ·0008 | − ·0043 | − 5·4 |
| 9 | + ·0044 | − ·0028 | − ·0024 | − ·0009 | + 0·4 |
| 10 | + ·0015 | − ·0026 | − ·0027 | + ·0012 | − 0·4 |

The graph of this curve is given in Fig. 1 (p. 323). The curve is drawn to scale but the distance to each finite mass is merely indicated, being unity or ten times the distance of the initial displacement of the infinitesimal body from the point I. The motion in this orbit is clockwise as indicated by the arrow.

If the infinitesimal body is projected from the negative $x$-axis at the initial distance 0·1, *i.e.* if $\epsilon = -0·1$, then the equations which define its path are

$$\left.\begin{aligned} x &= e^{-at}(-0·0996\cos\beta t - 0·0624\sin\beta t) \\ &\quad + e^{-2at}(0·0027 - 0·0031\cos 2\beta t - 0·0039\sin 2\beta t) + \ldots, \\ y &= e^{-at}(0·0067 + 0·1000\sin\beta t) \\ &\quad + e^{-2at}(-0·0068 + 0·0001\cos 2\beta t + 0·0012\sin 2\beta t) + \ldots, \\ x' &= e^{-at}(-0·0126\cos\beta t + 0·1030\sin\beta t) \\ &\quad + e^{-2at}(-0·0020 - 0·0039\cos 2\beta t + 0·0078\sin 2\beta t) + \ldots, \\ y' &= e^{-at}(-0·0775\cos\beta t - 0·0428\sin\beta t) \\ &\quad + e^{-2at}(0·0051 + 0·0018\cos 2\beta t - 0·0011\sin 2\beta t) + \ldots. \end{aligned}\right\} \quad \ldots\ldots\ldots\ldots(36)$$

By examining the value of $\dfrac{d^2y}{dx^2}$ we find that the curve represented by (36) is concave to the point I, as in the previous case. Table II contains a list of the coordinates and their derivatives for the same values of $t$ as in Table I.

TABLE II.

$\mu = 0·1 \qquad \epsilon = -0·1$

| $t$ | $x$ | $y$ | $x'$ | $y'$ | $\dfrac{dy}{dx} = \dfrac{y'}{x'}$ |
|---|---|---|---|---|---|
| 0 | − ·1 | 0 | − ·0185 | + ·0844 | − 4·7 |
| ·1 | − ·1015 | + ·0080 | − ·0086 | + ·0777 | − 9·0 |
| ·2 | − ·1017 | + ·0154 | + ·0008 | + ·0712 | + 8·9 |
| ·3 | − ·1012 | + ·0222 | + ·0095 | + ·0632 | + 6·7 |
| ·4 | − ·0998 | + ·0282 | + ·0172 | + ·0571 | + 3·3 |
| ·5 | − ·0980 | + ·0334 | + ·0243 | + ·0491 | + 2·0 |
| 1 | − ·0791 | + ·0500 | + ·0476 | + ·0180 | + 0·4 |
| 1·5 | − ·0531 | + ·0527 | + ·0543 | − ·0058 | − 0·1 |
| 2 | − ·0269 | + ·0457 | + ·0490 | − ·0205 | − 0·4 |
| 2·5 | − ·0049 | + ·0333 | + ·0379 | − ·0307 | − 0·8 |
| 3 | + ·0108 | + ·0196 | + ·0246 | − ·0274 | − 1·1 |
| 3·5 | + ·0199 | + ·0068 | + ·0118 | − ·0228 | − 1·9 |
| 4 | + ·0230 | − ·0030 | + ·0014 | − ·0165 | − 1·2 |
| 4·5 | + ·0218 | − ·0096 | − ·0064 | − ·0092 | + 1·4 |
| 5 | + ·0172 | − ·0124 | − ·0108 | − ·0027 | + 0·3 |
| 6 | + ·0057 | − ·0104 | − ·0107 | + ·0052 | − 0·5 |
| 7 | − ·0027 | − ·0043 | − ·0054 | + ·0064 | − 1·2 |
| 8 | − ·0053 | − ·0003 | − ·00003 | + ·0036 | − 1·20 |
| 9 | − ·0038 | + ·0028 | + ·0025 | + ·0004 | + 1·6 |
| 10 | − ·0018 | + ·0024 | + ·0022 | − ·0013 | − 0·6 |

The graph of this curve is given in Fig. 2. It is also to scale with the distances to the finite masses indicated as in Fig. 1. The curve is not only differently orientated from that in Fig. 1, but is of slightly different shape. The direction of motion is clockwise as in the preceding curve.

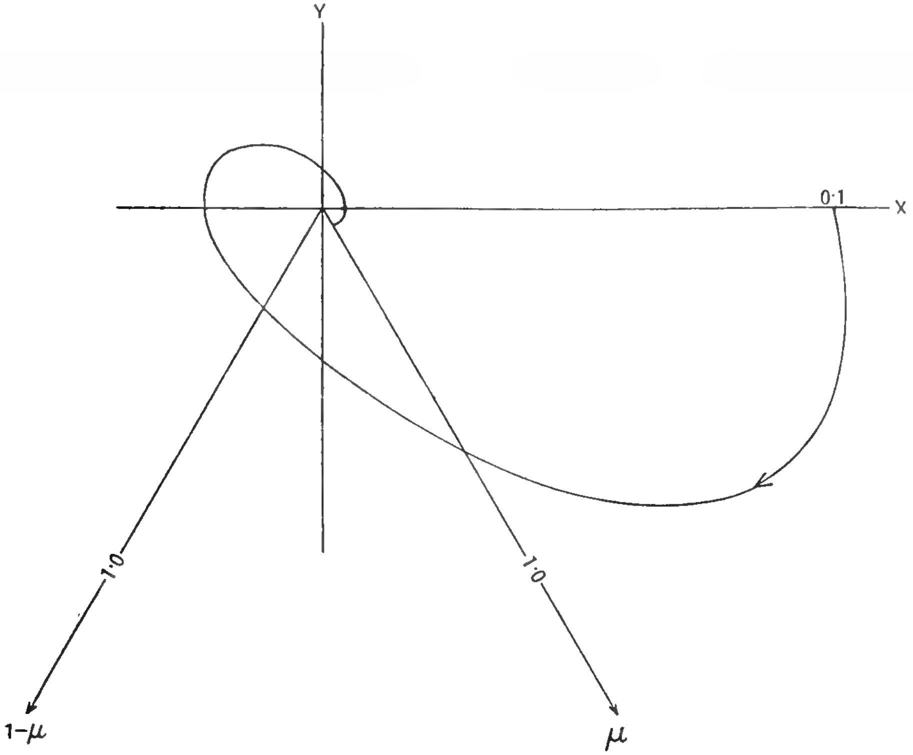



Fig. 1. The Orbit Asymptotic to the Point I for $\mu=0{\cdot}1$, $\epsilon=0{\cdot}1$.

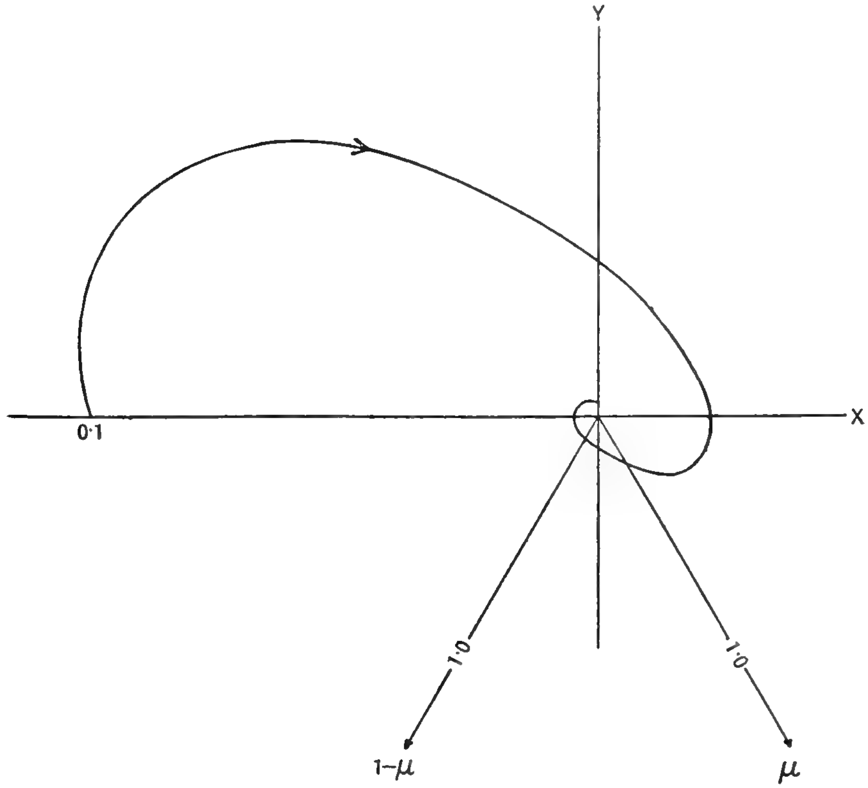



Fig. 2. The Orbit Asymptotic to the Point I for $\mu=0{\cdot}1$, $\epsilon=-0{\cdot}1$.

## PART II.

### THREE-DIMENSIONAL ASYMPTOTIC ORBITS.

6. THE DIFFERENTIAL EQUATIONS.

In the discussion of the three-dimensional asymptotic orbits, the same notation and units are chosen as in § 2, and the differential equations which define the motion of the infinitesimal body are the same as (1). Besides transferring the origin to the point I by equations (2) and introducing the parameter $\epsilon$ by (3), the independent variable is transformed by the substitution

$$t - t_0 = (1 + \delta)\,\tau, \qquad \text{(37)}$$

where $\delta$ is a function of $\epsilon$ so determined that the solutions of (1) shall be periodic with the period $2\pi$ in $\tau$. When the transformations (2), (3), and (37) have been made in (1), the differential equations of motion become

$$\left.\begin{aligned} \ddot{x} - 2(1+\delta)\,\dot{y} &= (1+\delta)^2\,[X_1 + X_2\epsilon + X_3\epsilon^2 + \ldots],\\ \ddot{y} + 2(1+\delta)\,\dot{x} &= (1+\delta)^2\,[Y_1 + Y_2\epsilon + Y_3\epsilon^2 + \ldots],\\ \ddot{z} + (1+\delta)^2\,z &= (1+\delta)^2\,[Z_2\epsilon + Z_3\epsilon^2 + \ldots], \end{aligned}\right\} \qquad \text{(38)}$$

where the dots denote differentiation with respect to $\tau$; and $X_n$, $Y_n$, and $Z_n$ are the same polynomials as in (5). The periodic solutions of these equations, in terms of the variables $x$, $y$, and $z$ of equations (2), are*

$$\left.\begin{aligned} \bar{\bar{x}} &= \epsilon x_0 \equiv \sum_{n=1}^{\infty}\sum_{k=0}^{n}[a_{2k}^{(2n)}\cos 2k\tau + \sqrt{3}\,b_{2k}^{(2n)}\sin 2k\tau]\,\epsilon^{2n},\\ y &= \epsilon y_0 \equiv \sum_{n=1}^{\infty}\sum_{k=0}^{n}[\sqrt{3}\,c_{2k}^{(2n)}\cos 2k\tau + d_{2k}^{(2n)}\sin 2k\tau]\,\epsilon^{2n},\\ \bar{\bar{z}} &= \epsilon z_0 \equiv \sum_{n=0}^{\infty}\sum_{k=0}^{\infty}[\sqrt{3}\,g_{2k+1}^{(2n+1)}\cos(2k+1)\,\tau + h_{2k+1}^{(2n+1)}\sin(2k+1)\,\tau]\,\epsilon^{2n+1},\\ \delta &= \delta_2\epsilon^2 + \delta_4\epsilon^4 + \ldots, \end{aligned}\right\} \qquad \text{(39)}$$

where the various coefficients of the cosines and sines, and also $\delta_2$, $\delta_4$, ..., are rational functions of $\mu$. It is not necessary in these periodic solutions to restrict $\mu$ to be less than $\mu_0 = \cdot 0385\ldots$, as in Part I. The radical $\sqrt{3}$ occurs in the solutions (39) only where indicated. The initial conditions are chosen so that

$$z_0(0) = 0, \quad \dot{z}_0(0) = 1. \qquad \text{(40)}$$

The numerical values of the coefficients in (39), in so far as they are given in "Oscillating Satellite", are

$$\left.\begin{aligned} & a_0^{(2)} = b_0^{(2)} = d_0^{(2)} = g_1^{(1)} = g_1^{(3)} = g_3^{(3)} = 0,\\ & a_2^{(2)} = \frac{8(1-2\mu)}{73 - 9(1-2\mu)^2}, \quad b_2^{(2)} = \frac{8}{73 - 9(1-2\mu)^2},\\ & c_0^{(2)} = -\tfrac{1}{6}, \quad c_2^{(2)} = \frac{19 - 3(1-2\mu)^2}{2\,[73 - 9(1-2\mu)^2]}, \quad d_2^{(2)} = -\frac{8(1-2\mu)}{73 - 9(1-2\mu)^2},\\ & h_1^{(3)} = -\frac{9\mu(1-\mu)}{73 - 9(1-2\mu)^2}, \quad h_3^{(3)} = \frac{3\mu(1-\mu)}{73 - 9(1-2\mu)^2}, \quad \delta_2 = \frac{12\mu(1-\mu)}{73 - 9(1-2\mu)^2}. \end{aligned}\right\} \qquad \text{(41)}$$

* "Oscillating Satellite," § 164.

## 7. The Equations of Variation.

We propose to construct solutions of the differential equations (38) which are asymptotic to the periodic solutions $x_0$, $y_0$, and $z_0$ in (39). Let

$$x = x_0 + p, \quad y = y_0 + q, \quad z = z_0 + r \quad \text{.......(42)}$$

be substituted in (38), $p$, $q$, and $r$ denoting new dependent variables. As a result we have

$$\left.\begin{aligned} \ddot{p} - 2(1+\delta)\dot{q} + (1+\delta)^2[P_1 p + P_2 q + P_3 r] &= (1+\delta)^2 P, \\ \ddot{q} + 2(1+\delta)\dot{p} + (1+\delta)^2[Q_1 p + Q_2 q + Q_3 r] &= (1+\delta)^2 Q, \\ \ddot{r} + (1+\delta)^2[R_1 p + R_2 q + R_3 r] &= (1+\delta)^2 R, \end{aligned}\right\} \quad \text{.......(43)}$$

where

$$\left.\begin{aligned} P_1 &= -\tfrac{3}{4} - \tfrac{3}{8}[7(1-2\mu)x_0 + \sqrt{3}y_0]\epsilon + \ldots, \\ P_2 &= -\tfrac{3}{4}\sqrt{3}(1-2\mu) - \tfrac{3}{8}[\sqrt{3}x_0 - 11(1-2\mu)y_0]\epsilon + \ldots, \\ P_3 &= -\tfrac{3}{2}(1-2\mu)z_0\epsilon + \ldots, \\ Q_1 &= -\tfrac{3}{4}\sqrt{3}(1-2\mu) - \tfrac{3}{8}[\sqrt{3}x_0 - 11(1-2\mu)y_0]\epsilon + \ldots, \\ Q_2 &= -\tfrac{9}{4} - \tfrac{3}{8}[11(1-2\mu)x_0 + 3\sqrt{3}y_0]\epsilon + \ldots, \\ Q_3 &= -\frac{3\sqrt{3}}{2}z_0\epsilon + \ldots, \\ R_1 &= -\tfrac{3}{2}(1-2\mu)z_0\epsilon + \ldots, \\ R_2 &= -\frac{3\sqrt{3}}{2}z_0\epsilon + \ldots, \\ R_3 &= 1 - \tfrac{3}{2}[(1-2\mu)x_0 + \sqrt{3}y_0]\epsilon + \ldots, \\ P &= +\tfrac{3}{16}[7(1-2\mu)p^2 + 2\sqrt{3}pq - 11(1-2\mu)q^2 + 4(1-2\mu)r^2]\epsilon + \ldots, \\ Q &= -\tfrac{3}{16}[\sqrt{3}p^2 + 22(1-2\mu)pq + 3\sqrt{3}q^2 - 4\sqrt{3}r^2]\epsilon + \ldots, \\ R &= +\tfrac{3}{2}[(1-2\mu)pr + \sqrt{3}qr]\epsilon + \ldots. \end{aligned}\right\} \quad \text{......(44)}$$

If the right members of (43) are neglected, we obtain

$$\left.\begin{aligned} \ddot{p} - 2(1+\delta)\dot{q} + (1+\delta)^2[P_1 p + P_2 q + P_3 r] &= 0, \\ \ddot{q} + 2(1+\delta)\dot{p} + (1+\delta)^2[Q_1 p + Q_2 q + Q_3 r] &= 0, \\ \ddot{r} + (1+\delta)^2[R_1 p + R_2 q + R_3 r] &= 0, \end{aligned}\right\} \quad \text{.......(45)}$$

which are the *equations of variation.*

## 8. The Solutions of the Equations of Variation.

The equations of variation are linear differential equations having periodic coefficients, the period being $2\pi$ in $\tau$. Such equations were first discussed by Hill* in 1877 in his celebrated memoir on the lunar theory. Since that time, these equations have been discussed extensively by Poincaré and many other prominent mathematicians†. The method which we shall adopt in constructing the solutions of (45) is the one developed by Moulton and Macmillan‡. This method of construction is essentially one of undetermined coefficients.

* The *Collected Works* of G. W. Hill, vol. I. p. 243: *Acta Mathematica*, vol. VIII. pp. 1–36.

† A very complete list of references to the literature of these differential equations is given by Baker on p. 134 of his memoir "On Certain Linear Differential Equations of Astronomical Interest," *Philosophical Transactions of the Royal Society of London*, Series A, vol. 216, pp. 129–186.

‡ "On the Solutions of Certain Types of Linear Differential Equations with Periodic Coefficients," *American Journal of Mathematics*, vol. XXXIII. No. 1 (1911).

The differential equations (45) are simultaneous and must be considered together. The general form of the solutions, first given by Floquet*, is

$$p = e^{\lambda\tau} u, \quad q = e^{\lambda\tau} v, \quad r = e^{\lambda\tau} w, \quad \text{.......(46)}$$

where an existence proof, following the method of Moulton and Macmillan, would show that $\lambda$, $u$, and $v$ are power series in $\epsilon^2$ and that $w$ is a power series in odd powers of $\epsilon$. Let

$$\lambda = \lambda_0 + \lambda_2 \epsilon^2 + \dots \quad \text{.......(47)}$$

There are six values of $\lambda_0$ altogether, but two of them are known to be zero†, since the generating solutions contain two arbitrary constants $t_0$ and $\epsilon$. Considering first the exponents which are not identically zero, we observe from (7), Part I, that the values of $\lambda_0^2$ are

$$\lambda_0^2 = \frac{-1 \pm \sqrt{1 - 27\mu(1-\mu)}}{2}.$$

Since at least one value of $\lambda_0$ must be real or complex in order to construct asymptotic solutions, $\mu$ must be restricted to be greater than $\mu_0 = \cdot 0385\dots$, as in Part I, in which case the four values of $\lambda_0$ are all complex and differ in pairs only in sign. Let these values of $\lambda_0$ be denoted by

$$\sigma_0, \; -\sigma_0, \; \bar{\sigma}_0, \text{ and } -\bar{\sigma}_0.$$

These are the same values as $\sigma$ and $\bar{\sigma}$ in Part I.

Now let

$$\left.\begin{aligned} u &= u^{(0)} + u^{(2)}\epsilon^2 + \dots + u^{(2n)}\epsilon^{2n} + \dots, \\ v &= v^{(0)} + v^{(2)}\epsilon^2 + \dots + v^{(2n)}\epsilon^{2n} + \dots, \\ w &= w^{(1)}\epsilon + w^{(3)}\epsilon^3 + \dots + w^{(2n-1)}\epsilon^{2n-1} + \dots, \end{aligned}\right\} \quad \text{.......(48)}$$

where the various coefficients can be determined by a proper choice of $\lambda$ so that they shall be periodic with the period $2\pi$ in $\tau$. Since these solutions are later multiplied by arbitrary constants, we may assume, without loss of generality, that $u(0) = 1$, from which it follows that

$$u^{(0)}(0) = 1, \quad u^{(2j)}(0) = 0, \quad (j = 1, \dots \infty). \quad \text{.......(49)}$$

The initial values of $v$ and $w$ will then be determined from the differential equations.

On substituting (48) and (46) in (45), putting

$$\lambda = \sigma \equiv \sigma_0 + \sigma_2\epsilon^2 + \dots, \quad \text{.......(50)}$$

and equating to zero the coefficients of the various powers of $\epsilon$, we obtain sets of differential equations which define the various $u^{(2j)}$, $v^{(2j)}$, and $w^{(2j+1)}$. The constants of integration that arise from the solutions of these sets of differential equations are uniquely determined from the conditions (49). The solutions themselves can be made periodic by a proper choice of the various $\sigma_{2j}$ in (50). One set of solutions is thus found to be

$$p = e^{\sigma\tau} u_1, \quad q = e^{\sigma\tau} v_1, \quad r = e^{\sigma\tau} w_1, \quad \text{.......(51)}$$

where

$$\left.\begin{aligned} u_1 &= \sum_{j=0}^{\infty} u_1^{(2j)}\epsilon^{2j} = \sum_{j=0}^{\infty}\sum_{k=0}^{j} [F_{2k}^{(2j)} \cos 2k\tau + G_{2k}^{(2j)} \sin 2k\tau]\,\epsilon^{2j}, \\ v_1 &= \sum_{j=0}^{\infty} v_1^{(2j)}\epsilon^{2j} = \sum_{j=0}^{\infty}\sum_{k=0}^{j} [H_{2k}^{(2j)} \cos 2k\tau + K_{2k}^{(2j)} \sin 2k\tau]\,\epsilon^{2j}, \\ w_1 &= \sum_{j=0}^{\infty} w_1^{(2j+1)}\epsilon^{2j+1} = \sum_{j=0}^{\infty}\sum_{k=0}^{j} [L_{2k+1}^{(2j+1)} \cos(2k+1)\tau + M_{2k+1}^{(2j+1)} \sin(2k+1)\tau]\,\epsilon^{2j+1}, \\ \sigma &= \alpha + i\beta = \sum_{j=0}^{\infty} [\alpha_{2j} + i\beta_{2j}]\,\epsilon^{2j}. \end{aligned}\right\} \quad \text{...(52)}$$

* *Annales de l'École Normale Supérieure*, 1883–1884.

† Poincaré, *Mécanique Céleste*, vol. I. chap. IV.

The coefficients of the cosines and sines in (52) are complex numbers and contain some terms having $\sqrt{3}$ as a factor. The $\alpha_{2j}$ and $\beta_{2j}$ are real constants. The values of $\sigma_0$, $u_1^{(0)}$, $v_1^{(0)}$, and $w_1^{(0)}$ are

$$\sigma_0 = \alpha_0 + i\beta_0 = \tfrac{1}{2}\left[\{\sqrt{27\mu(1-\mu)} - 1\}^{\frac{1}{2}} + i\,\{\sqrt{27\mu(1-\mu)} + 1\}^{\frac{1}{2}}\right],$$

$$u_1^{(0)} = 1, \quad v_1^{(0)} = \frac{\sigma_0^2 - \frac{3}{4}}{2\sigma_0 + \frac{3}{4}\sqrt{3}\,(1-2\mu)} = \frac{\frac{3}{4}\sqrt{3}\,(1-2\mu) - 2\sigma_0}{\sigma_0^2 - \frac{9}{4}},$$

$$w_1^{(0)} = \frac{\frac{3}{2}(1-2\mu) + \dfrac{3\sqrt{3}}{2} v_1^{(0)}}{\sigma_0^2 + 4}\left[\sin\tau - \frac{2}{\sigma_0}\cos\tau\right].$$

By putting $\lambda_0$ equal to $-\sigma_0$, $\bar{\sigma}_0$, and $-\bar{\sigma}_0$, the corresponding solutions of (45) could be constructed in the same way as the solutions (51) were obtained. It is not necessary to repeat this construction, however, as we propose to show that these solutions can be obtained from (51) by changing the signs of $\sqrt{3}$, $\tau$, $p$, $q$, and $i$ in (51).

We shall first show that the differential equations (43) are unchanged if the signs of $\sqrt{3}$, $\tau$, $q$, and $r$ are changed, and, obviously, (45) will remain the same under these changes. The discussion is made for the equations (43), instead of for (45), as this property of (43) is used later in § 9 (B).

The function $U$ of the original differential equations (1) is even in $\eta$ and $\zeta$. Hence $\dfrac{\partial U}{\partial \xi}$ is likewise even in $\eta$ and $\zeta$, while $\dfrac{\partial U}{\partial \eta}$ is odd in $\eta$ but even in $\zeta$, and $\dfrac{\partial U}{\partial \zeta}$ is even in $\eta$ but odd in $\zeta$. Further, $\dfrac{\partial U}{\partial \eta}$ carries the factor $\eta$, and $\dfrac{\partial U}{\partial \zeta}$ carries the factor $\zeta$. When equations (1) are transformed by (2), (3), (37), and (42), that is, when the substitutions

$$\xi = \tfrac{1}{2} - \mu + \epsilon(x_0 + p), \quad \eta = +\tfrac{1}{2}\sqrt{3} + \epsilon(y_0 + q), \quad \zeta = \epsilon(z_0 + r), \quad t - t_0 = (1+\delta)\tau \qquad \ldots(53)$$

are made, we obtain equations (43). The above transformation of the independent variable, viz. $t - t_0 = (1+\delta)\tau$, does not affect the parity of the equivalent expressions for $\xi$, $\eta$, and $\zeta$ in (43). Since the substitutions for the dependent variables are linear in (53), the equivalent expressions for $\xi$, $\eta$, and $\zeta$ enter (43) with the same parity as $\xi$, $\eta$, and $\zeta$ enter (1).

Now the terms of (43) which carry the factor $(1+\delta)^2$ arise from the right members of (1). Consider the expression

$$(1+\delta)^2\,[P_1 p + P_2 q + P_3 r - P]. \qquad \ldots\ldots\ldots\ldots(43, 1)$$

Since $\dfrac{\partial U}{\partial \xi}$ in (1) is even in $\eta$ and $\zeta$, (43, 1) is even in $(\sqrt{3}, y_0, q)$, considered together, and even also in $(z_0, r)$, taken together. Hence if we change the signs of $y_0$, $z_0$, $q$, and $r$ in (43, 1), and also of $\sqrt{3}$ in so far as it enters explicitly, the expression (43, 1) will remain unchanged.

The expression

$$(1+\delta)^2\,[Q_1 p + Q_2 q + Q_3 r - Q], \qquad \ldots\ldots\ldots\ldots(43, 2)$$

from the second equation of (43), carries a factor which is odd in $(\sqrt{3}, y_0, q)$, considered together, since $\dfrac{\partial U}{\partial \eta}$ carries the factor $\eta$ and is odd in $\eta$. Further, this expression is even in $(z_0, r)$, taken together. Therefore if the signs of $y_0$, $z_0$, $q$, and $r$, and also of $\sqrt{3}$ in so far as it enters explicitly, are changed in (43, 2) the expression changes sign.

The corresponding expression from the last equation of (43) is

$$(1+\delta)^2[R_1 p + R_2 q + R_3 r - R]. \qquad (43, 3)$$

Since $\frac{\partial U}{\partial \zeta}$ carries the factor $\zeta$ and is odd in $\zeta$, (43, 3) will carry a factor in $(z_0, r)$ to odd degrees, taken together. As $\frac{\partial U}{\partial \zeta}$ is even in $\eta$, (43, 3) is even in $(\sqrt{3}, y_0, q)$, considered together. Therefore a change in the signs of $y_0$, $z_0$, $q$, and $r$, and also of $\sqrt{3}$ in so far as it occurs explicitly, will change the sign of (43, 3).

Considering the solutions (39), we observe that

$$x_0(+\sqrt{3}, +\tau) = x_0(-\sqrt{3}, -\tau), \quad y_0(+\sqrt{3}, +\tau) = -y_0(-\sqrt{3}, -\tau), \quad z_0(+\sqrt{3}, +\tau) = -z_0(-\sqrt{3}, -\tau).$$

Hence changing the signs of $y_0$ and $z_0$ is equivalent to changing the sign of $\tau$, and also of $\sqrt{3}$ where it occurs implicitly. Therefore changing the signs of $\tau$, $q$, $r$, and of $\sqrt{3}$ where it occurs explicitly and also implicitly, is equivalent to changing the signs of $y_0$, $z_0$, $q$, $r$, and of $\sqrt{3}$ where it occurs explicitly. Consequently, if we change the signs of $\tau$, $q$, $r$, and of $\sqrt{3}$ both implicitly and explicitly, the expression (43, 1) remains unchanged, while (43, 2) and (43, 3) both change signs. On making the same changes in the remaining terms of (43), we observe that the differential equations remain unchanged. Obviously, the same property holds for (43) if we neglect the right members, that is, for equations (45).

Let us proceed now to the determination of the remaining solutions of (45). Since these differential equations are unchanged by changing the signs of $\tau$, $q$, $r$, and $\sqrt{3}$, we shall obtain another set of solutions if we make the corresponding changes in the solutions (51). Let this set be denoted by

$$p = e^{-\sigma\tau} u_2, \quad q = -e^{-\sigma\tau} v_2, \quad r = -e^{-\sigma\tau} w_2, \qquad (54)$$

where $u_2$, $v_2$, and $w_2$ differ respectively from $u_1$, $v_1$, and $w_1$ only in the signs of $\sqrt{3}$ and $\tau$. Thus

$$\left.\begin{aligned} u_2 &= \sum_{j=0}^{\infty} u_2^{(2j)} \epsilon^{2j} \equiv \sum_{j=0}^{\infty} \sum_{k=0}^{j} [\hat{F}_{2k}^{(2j)} \cos 2k\tau - \hat{G}_{2k}^{(2j)} \sin 2k\tau]\, \epsilon^{2j}, \\ v_2 &= \sum_{j=0}^{\infty} v_2^{(2j)} \epsilon^{2j} \equiv \sum_{j=0}^{\infty} \sum_{k=0}^{j} [\hat{H}_{2k}^{(2j)} \cos 2k\tau - \hat{K}_{2k}^{(2j)} \sin 2k\tau]\, \epsilon^{2j}, \\ w_2 &= \sum_{j=0}^{\infty} w_2^{(2j+1)} \epsilon^{2j+1} \equiv \sum_{j=0}^{\infty} \sum_{k=0}^{j} [\hat{L}_{2k+1}^{(2j+1)} \cos (2k+1)\tau - \hat{M}_{2k+1}^{(2j+1)} \sin (2k+1)\tau]\, \epsilon^{2j+1}, \end{aligned}\right\} \qquad (55)$$

where the circumflex (^) denotes that two constants $F$ and $\hat{F}$ differ only in the sign of $\sqrt{3}$.

Since the differential equations (45) are independent of $i$, a change in the sign of $i$ will leave the differential equations unaltered; and if we change the sign of $i$ in (51) and (54) we obtain the additional solutions

$$p = e^{\bar{\sigma}\tau} \bar{u}_1, \quad q = e^{\bar{\sigma}\tau} \bar{v}_1, \quad r = e^{\bar{\sigma}\tau} \bar{w}_1, \qquad (56)$$

and

$$p = e^{-\bar{\sigma}\tau} \bar{u}_2, \quad q = -e^{-\bar{\sigma}\tau} \bar{v}_2, \quad r = -e^{-\bar{\sigma}\tau} \bar{w}_2, \qquad (57)$$

respectively.

The solutions (51), (54), (56), and (57) are those with characteristic exponents different from zero. We shall now derive the solutions of (45) which have zero characteristic exponents.

It has been shown by Poincaré* that if the generating solutions (39) contain an arbitrary constant which does not occur explicitly in the original differential equations (1), then a solution

* Poincaré, *Mécanique Céleste*, vol. I. chap. IV.

of the equations of variation can be obtained by differentiating the generating solutions with respect to this constant. The generating solutions (39) contain two such arbitrary constants, $t_0$ and $\epsilon$, and therefore the two remaining solutions of (45) can be obtained by differentiating (39) with respect to these constants.

Consider first the constant $t_0$. It enters (39) implicitly through $\tau$. Thus one set of solutions is

$$p = \frac{\partial(\epsilon x_0)}{\partial t_0} = -\frac{1}{1+\delta}\frac{\partial(\epsilon x_0)}{\partial \tau} \equiv -\frac{1}{1+\delta} u_3,$$

$$q = \frac{\partial(\epsilon y_0)}{\partial t_0} = -\frac{1}{1+\delta}\frac{\partial(\epsilon y_0)}{\partial \tau} \equiv -\frac{1}{1+\delta} v_3,$$

$$r = \frac{\partial(\epsilon z_0)}{\partial t_0} = -\frac{1}{1+\delta}\frac{\partial(\epsilon z_0)}{\partial \tau} \equiv -\frac{1}{1+\delta} w_3.$$

Since these solutions are later multiplied by arbitrary constants, the constant multiplier $-\frac{1}{1+\delta}$ may be absorbed, and we may take

$$p = u_3, \quad q = v_3, \quad r = w_3 \quad \text{.................................(58)}$$

as the solutions. The differentiations give

$$u_3 = \frac{\partial(\epsilon x_0)}{\partial \tau} = \frac{16}{73 - 9(1-2\mu)^2}[\sqrt{3}\cos 2\tau - (1-2\mu)\sin 2\tau]\epsilon^2 + \ldots,$$

$$v_3 = \frac{\partial(\epsilon y_0)}{\partial \tau} = \frac{-1}{73 - 9(1-2\mu)^2}[16(1-2\mu)\cos\tau + \{19 - 3(1-2\mu)^2\}\sqrt{3}\sin 2\tau]\epsilon^2 + \ldots.$$

$$w_3 = \frac{\partial(\epsilon z_0)}{\partial \tau} = [\cos\tau]\epsilon + \frac{9\mu(1-\mu)}{73 - 9(1-2\mu)^2}[-\cos\tau + \cos 3\tau]\epsilon^3 + \ldots.$$

On differentiating (39) with respect to $\epsilon$ we obtain the solutions

$$p = \frac{\partial(\epsilon x_0)}{\partial \epsilon}, \qquad q = \frac{\partial(\epsilon y_0)}{\partial \epsilon}, \qquad r = \frac{\partial(\epsilon z_0)}{\partial \epsilon}.$$

Since $\epsilon$ enters (39) explicitly and also implicitly through $\tau$, we have

$$\left.\begin{aligned} p &= \frac{\partial(\epsilon x_0)}{\partial \epsilon} = \left(\frac{\partial(\epsilon x_0)}{\partial \epsilon}\right) + \frac{\partial(\epsilon x_0)}{\partial \tau}\frac{\partial \tau}{\partial \delta}\frac{\partial \delta}{\partial \epsilon} \equiv u_4 + K\tau u_3, \\ q &= \frac{\partial(\epsilon y_0)}{\partial \epsilon} = \left(\frac{\partial(\epsilon y_0)}{\partial \epsilon}\right) + \frac{\partial(\epsilon y_0)}{\partial \tau}\frac{\partial \tau}{\partial \delta}\frac{\partial \delta}{\partial \epsilon} \equiv v_4 + K\tau v_3, \\ r &= \frac{\partial(\epsilon z_0)}{\partial \epsilon} = \left(\frac{\partial(\epsilon z_0)}{\partial \epsilon}\right) + \frac{\partial(\epsilon z_0)}{\partial \tau}\frac{\partial \tau}{\partial \delta}\frac{\partial \delta}{\partial \epsilon} \equiv w_4 + K\tau w_3, \end{aligned}\right\} \quad \text{.....................(59)}$$

where the parentheses ( ) about the partial derivatives denote that the differentiation is performed only in so far as $\epsilon$ occurs explicitly. The differentiations give

$$u_4 = \left(\frac{\partial(\epsilon x_0)}{\partial \epsilon}\right) = \frac{16}{73 - 9(1-2\mu)^2}[(1-2\mu)\cos 2\tau + \sqrt{3}\sin 2\tau]\epsilon + \ldots,$$

$$v_4 = \left(\frac{\partial(\epsilon y_0)}{\partial \epsilon}\right) = \left[-\frac{\sqrt{3}}{2} + \frac{\sqrt{3}\{19 - 3(1-2\mu)^2\}}{73 - 9(1-2\mu)^2}\cos 2\tau - \frac{16(1-2\mu)}{73 - 9(1-2\mu)^2}\sin 2\tau\right]\epsilon + \ldots,$$

$$w_4 = \left(\frac{\partial(\epsilon z_0)}{\partial \epsilon}\right) = \sin\tau - \frac{9\mu(1-\mu)}{73 - 9(1-2\mu)^2}[3\sin\tau - \sin 3\tau]\epsilon^2 + \ldots,$$

$$K = -\frac{1}{1+\delta}\frac{\partial \delta}{\partial \epsilon} = -\frac{1}{1+\delta}[2\delta_2\epsilon + 4\delta_4\epsilon^3 + \ldots].$$

This completes the integration of the equations of variation.

It remains now to show that the six solutions (51), (54), (56), (57), (58), and (59) constitute a fundamental set. The criterion for a fundamental set is that the determinant of these solutions and their first derivatives with respect to $\tau$ shall be different from zero. This determinant is

$$\Delta = \begin{vmatrix} e^{\sigma\tau}u_1, & e^{-\sigma\tau}u_2, & e^{\bar{\sigma}\tau}\bar{u}_1, & e^{-\bar{\sigma}\tau}\bar{u}_2, & u_3, & u_4 + K\tau u_3 \\ e^{\sigma\tau}(\sigma u_1 + \dot{u}_1), & -e^{-\sigma\tau}(\sigma u_2 - \dot{u}_2), & e^{\bar{\sigma}\tau}(\bar{\sigma}\bar{u}_1 + \dot{\bar{u}}_1), & -e^{-\bar{\sigma}\tau}(\bar{\sigma}\bar{u}_2 - \dot{\bar{u}}_2), & \dot{u}_3, & \dot{u}_4 + K(u_4 + \tau\dot{u}_3) \\ e^{\sigma\tau}v_1, & -e^{-\sigma\tau}v_2, & e^{\bar{\sigma}\tau}\bar{v}_1, & -e^{-\bar{\sigma}\tau}\bar{v}_2, & v_3, & v_4 + K\tau v_3 \\ e^{\sigma\tau}(\sigma v_1 + \dot{v}_1), & e^{-\sigma\tau}(\sigma v_2 - \dot{v}_2), & e^{\bar{\sigma}\tau}(\bar{\sigma}\bar{v}_1 + \dot{\bar{v}}_1), & e^{-\bar{\sigma}\tau}(\bar{\sigma}\bar{v}_2 - \dot{\bar{v}}_2), & \dot{v}_3, & \dot{v}_4 + K(v_4 + \tau\dot{v}_3) \\ e^{\sigma\tau}w_1, & -e^{-\sigma\tau}w_2, & e^{\bar{\sigma}\tau}\bar{w}_1, & -e^{-\bar{\sigma}\tau}\bar{w}_2, & w_3, & w_4 + K\tau w_3 \\ e^{\sigma\tau}(\sigma w_1 + \dot{w}_1), & e^{-\sigma\tau}(\sigma w_2 - \dot{w}_2), & e^{\bar{\sigma}\tau}(\bar{\sigma}\bar{w}_1 + \dot{\bar{w}}_1), & e^{-\bar{\sigma}\tau}(\bar{\sigma}\bar{w}_2 - \dot{\bar{w}}_2), & \dot{w}_3, & \dot{w}_4 + K(w_4 + \tau\dot{w}_3) \end{vmatrix}$$

It is a constant * for all values of $\tau$, and therefore the computation will be simplified by putting $\tau = 0$. Thus we find that

$$\Delta = \frac{-24\epsilon\beta_0^2(\alpha_0 + i\beta_0)\left\{\tfrac{3}{4}(1-2\mu)(3\alpha_0^2 - \beta_0^2 - \tfrac{9}{4}) - \dfrac{4}{\sqrt{3}}\alpha_0(\alpha_0^2 + \beta_0^2)\right\}\left[3\alpha_0(1-2\mu) - \tfrac{4}{3}(\tfrac{9}{4} + \alpha_0^2 + \beta_0^2)\right]}{(\alpha_0^2 + \beta_0^2)^2 - \tfrac{9}{4}(\alpha_0^2 - \beta_0^2) + \tfrac{81}{16}} \quad \ldots\ldots(60)$$

$$+ \text{ terms of higher degree in } \epsilon.$$

Since $\alpha_0 = \tfrac{1}{2}[\sqrt{27\mu(1-\mu)} - 1]^{\frac{1}{2}}, \quad \beta_0 = \tfrac{1}{2}[\sqrt{27\mu(1-\mu)} + 1]^{\frac{1}{2}}$,

and since $\mu$ is restricted to the interval $\mu_0 \leqq \mu \leqq \tfrac{1}{2}$, the product $\beta_0^2(\alpha_0 + i\beta_0)$ in (60) is different from zero for all values of $\mu$ in the above interval. The determinant $\Delta$ can vanish then only with either factor { } or [ ]. On equating each of these factors to zero, rationalizing and simplifying, we obtain, respectively, the equations

$$\mu^8 - 4\mu^7 + 4{\cdot}888\mu^6 - 0{\cdot}663\mu^5 - 5{\cdot}878\mu^4 + 3{\cdot}809\mu^3 - 0{\cdot}388\mu^2 - 0{\cdot}953\mu + 0{\cdot}0153 = 0,$$
$$\mu^6 - 3\mu^5 + 0{\cdot}837\mu^4 + 3{\cdot}325\mu^3 - 1{\cdot}019\mu^2 - 1{\cdot}144\mu + 0{\cdot}057 = 0.$$

By applying Sturm's† theorem we find that neither of the above equations has a root for $\mu$ lying between $\mu_0$ and $\tfrac{1}{2}$. Hence, for $\epsilon$ different from zero but sufficiently small numerically, and for all values of $\mu$ between $\mu_0$ and $\tfrac{1}{2}$, the determinant $\Delta$ is different from zero. Thus the six solutions which have been determined constitute a fundamental set of solutions of the equations of variation, and the most general solutions of these equations are

$$\left.\begin{aligned} p &= N_1 e^{\sigma\tau}u_1 + N_2 e^{-\sigma\tau}u_2 + N_3 e^{\bar{\sigma}\tau}\bar{u}_1 + N_4 e^{-\bar{\sigma}\tau}\bar{u}_2 + N_5 u_3 + N_6(u_4 + K\tau u_3), \\ q &= N_1 e^{\sigma\tau}v_1 - N_2 e^{-\sigma\tau}v_2 + N_3 e^{\bar{\sigma}\tau}\bar{v}_1 - N_4 e^{-\bar{\sigma}\tau}\bar{v}_2 + N_5 v_3 + N_6(v_4 + K\tau v_3), \\ r &= N_1 e^{\sigma\tau}w_1 - N_2 e^{-\sigma\tau}w_2 + N_3 e^{\bar{\sigma}\tau}\bar{w}_1 - N_4 e^{-\bar{\sigma}\tau}\bar{w}_2 + N_5 w_3 + N_6(w_4 + K\tau w_3), \end{aligned}\right\} \quad \ldots\ldots(61)$$

where $N_1, \ldots, N_6$ are arbitrary constants. This completes the construction of the solutions of the equations of variation.

## 9. Construction of Asymptotic Solutions.

### (A) *Solutions in* $e^{-\alpha\tau}$.

In making the construction of the asymptotic solutions of (43), it is convenient to introduce another parameter $\gamma$ by the substitutions

$$\bar{p} = p\gamma, \quad \bar{q} = q\gamma, \quad \bar{r} = r\gamma. \quad \ldots\ldots(62)$$

* Moulton, *Periodic Orbits*, chap. I. § 18.

† Burnside and Panton, *Theory of Equations*, vol. I. p. 198.

where $\bar{\bar{p}}$, $\bar{\bar{q}}$, and $\bar{\bar{r}}$ are the new dependent variables. Let

$$\left.\begin{aligned}\bar{\bar{p}} &= p_0 + p_1\gamma + p_2\gamma^2 + \ldots,\\ \bar{\bar{q}} &= q_0 + q_1\gamma + q_2\gamma^2 + \ldots,\\ \bar{\bar{r}} &= r_0 + r_1\gamma + r_2\gamma^2 + \ldots.\end{aligned}\right\} \qquad (63)$$

When (62) and (63) are substituted in (43), the factor $\gamma$ will divide out of the resulting equations and we obtain differential equations which are to be satisfied identically in $\gamma$. By equating the coefficients of the same powers of $\gamma$ in these equations, which will be cited as (43′), we obtain sets of differential equations which determine the various $p_j$, $q_j$, and $r_j$ in (63). In order to obtain asymptotic solutions it is necessary to impose suitable conditions on the solutions of these equations.

According to Poincaré's definition, each term of an asymptotic solution must contain a factor of the form $e^{\lambda\tau}$, where $\lambda$ is real or complex. Obviously, the only exponents which enter into the integrations of (43) are those which arise from the solutions of the equations of variation. These exponents are $\pm\sigma$ and $\pm\bar{\sigma}$. Considering first the solutions of (43) which approach zero as $\tau$ approaches $+\infty$, we must impose the condition ($C_1$), that each term of the solutions must contain the factor $e^{-\sigma\tau}$ or $e^{-\bar{\sigma}\tau}$, as these are the only exponentials which have their real parts negative.

It is evident that at each step of the integration of (43′) two arbitrary constants will arise which are not determined by condition ($C_1$). These are the constants associated with $e^{-\sigma\tau}$ and $e^{-\bar{\sigma}\tau}$. In order that they may be uniquely determined we impose the conditions ($C_2$), that $\bar{\bar{p}}(0) = a$, $\bar{\bar{q}}(0) = 0$. As a consequence of these conditions we have from (63)

$$\left.\begin{aligned}&p_0(0) = a, \quad p_j(0) = 0, \quad (j = 1, 2, \ldots \infty),\\ &q_j(0) = 0, \quad (j = 0, 1, \ldots \infty).\end{aligned}\right\} \qquad (64)$$

Now consider the various coefficients of $\gamma$ in the equations (43′). The differential equations obtained from the terms which are independent of $\gamma$ are the same as the equations of variation except for the subscript 0. The solutions which satisfy ($C_1$) are therefore

$$\left.\begin{aligned}p_0 &= \phantom{-}N_2^{(0)}e^{-\sigma\tau}u_2 + N_4^{(0)}e^{-\bar{\sigma}\tau}\bar{u}_2,\\ q_0 &= -N_2^{(0)}e^{-\sigma\tau}v_2 - N_4^{(0)}e^{-\bar{\sigma}\tau}\bar{v}_2,\\ r_0 &= -N_2^{(0)}e^{-\sigma\tau}w_2 - N_4^{(0)}e^{-\bar{\sigma}\tau}\bar{w}_2,\end{aligned}\right\} \qquad (65)$$

where $N_2^{(0)}$ and $N_4^{(0)}$ are constants of integration. On imposing conditions ($C_2$), we have from (64)

$$N_2^{(0)} + N_4^{(0)} = a, \qquad N_2^{(0)}v_2(0) + N_4^{(0)}\bar{v}_2(0) = 0,$$

from which it follows that

$$N_2^{(0)} = \frac{a\bar{v}_2(0)}{\bar{v}_2(0) - v_2(0)}, \qquad N_4^{(0)} = \bar{N}_2^{(0)}.$$

Hence the solutions (65) may be written

$$\left.\begin{aligned}p_0 &= a\left[e^{-\sigma\tau}u_{10}^{(0)} + e^{-\bar{\sigma}\tau}u_{01}^{(0)}\right],\\ q_0 &= a\left[e^{-\sigma\tau}v_{10}^{(0)} + e^{-\bar{\sigma}\tau}v_{01}^{(0)}\right],\\ r_0 &= a\left[e^{-\sigma\tau}w_{10}^{(0)} + e^{-\bar{\sigma}\tau}w_{01}^{(0)}\right],\end{aligned}\right\} \qquad (66)$$

where $u_{10}^{(0)}$, $v_{10}^{(0)}$, and $w_{10}^{(0)}$ are similar in form to $u_1$, $v_1$, and $w_1$ respectively; and

$$u_{01}^{(0)} = \bar{u}_{10}^{(0)}, \quad v_{01}^{(0)} = \bar{v}_{10}^{(0)}, \quad w_{01}^{(0)} = \bar{w}_{10}^{(0)}.$$

The differential equations obtained by equating the coefficients of $j$ in (43′) have the same left members as the equations of variation except for the subscript 1 on $p$, $q$, and $r$. The right members, which we denote by $P^{(1)}$, $Q^{(1)}$, and $R^{(1)}$ respectively, have the forms

$$P^{(1)} = \epsilon a^2 [e^{-2\sigma\tau} U_{20}^{(1)} + e^{-(\sigma+\bar{\sigma})\tau} U_{11}^{(1)} + e^{-2\bar{\sigma}\tau} U_{02}^{(1)}],$$
$$Q^{(1)} = \epsilon a^2 [e^{-2\sigma\tau} V_{20}^{(1)} + e^{-(\sigma+\bar{\sigma})\tau} V_{11}^{(1)} + e^{-2\bar{\sigma}\tau} V_{02}^{(1)}],$$
$$R^{(1)} = \epsilon a^2 [e^{-2\sigma\tau} W_{20}^{(1)} + e^{-(\sigma+\bar{\sigma})\tau} W_{11}^{(1)} + e^{-2\bar{\sigma}\tau} W_{02}^{(1)}],$$

where $U_{20}^{(1)}$, $V_{20}^{(1)}$, and $W_{20}^{(1)}$ are similar to $u_1$, $v_1$, and $w_1$ respectively; and

$$U_{02}^{(1)} = \bar{U}_{20}^{(1)}, \quad V_{02}^{(1)} = \bar{V}_{20}^{(1)}, \quad W_{02}^{(1)} = \bar{W}_{20}^{(1)}.$$

The functions $U_{11}^{(1)}$, $V_{11}^{(1)}$, and $W_{11}^{(1)}$ are also similar to $u_1$, $v_1$, and $w_1$ respectively, except that the coefficients of the cosines and sines are real and not complex.

The complementary functions of the differential equations defined in the preceding paragraph are the same as (61), but we shall denote the constants of integration by $n_1^{(1)}, \ldots, n_6^{(1)}$ instead of $N_1, \ldots, N_6$, respectively. The particular integrals of these equations can be obtained by the method of the variation of parameters. According to this method we have

$$\left.\begin{aligned}
&\dot{n}_1^{(1)} e^{\sigma\tau} u_1 + \dot{n}_2^{(1)} e^{-\sigma\tau} u_2 + \dot{n}_3^{(1)} e^{\bar{\sigma}\tau} \bar{u}_1 + \dot{n}_4^{(1)} e^{-\bar{\sigma}\tau} \bar{u}_2 + \dot{n}_5^{(1)} u_3 + \dot{n}_6^{(1)} (u_4 + K\tau u_3) = 0,\\
&\dot{n}_1^{(1)} e^{\sigma\tau} (\sigma u_1 + \dot{u}_1) - \dot{n}_2^{(1)} e^{-\sigma\tau} (\sigma u_2 - \dot{u}_2) + \dot{n}_3^{(1)} e^{\bar{\sigma}\tau} (\bar{\sigma}\bar{u}_1 + \dot{\bar{u}}_1) - \dot{n}_4^{(1)} e^{-\bar{\sigma}\tau} (\bar{\sigma}\bar{u}_2 - \dot{\bar{u}}_2)\\
&\qquad + \dot{n}_5^{(1)} \dot{u}_3 + \dot{n}_6^{(1)} [\dot{u}_4 + K(u_4 + \tau\dot{u}_3)] = P^{(1)},\\
&\dot{n}_1^{(1)} e^{\sigma\tau} v_1 - \dot{n}_2^{(1)} e^{-\sigma\tau} v_2 + \dot{n}_3^{(1)} e^{\bar{\sigma}\tau} \bar{v}_1 - \dot{n}_4^{(1)} e^{-\bar{\sigma}\tau} \bar{v}_2 + \dot{n}_5^{(1)} v_3 + \dot{n}_6^{(1)} (v_4 + K\tau v_3) = 0,\\
&\dot{n}_1^{(1)} e^{\sigma\tau} (\sigma v_1 + \dot{v}_1) + \dot{n}_2^{(1)} e^{-\sigma\tau} (\sigma v_2 - \dot{v}_2) + \dot{n}_3^{(1)} e^{\bar{\sigma}\tau} (\sigma\bar{v}_1 + \dot{\bar{v}}_1) + \dot{n}_4^{(1)} e^{-\bar{\sigma}\tau} (\bar{\sigma}\bar{v}_2 - \dot{\bar{v}}_2)\\
&\qquad + \dot{n}_5^{(1)} \dot{v}_3 + \dot{n}_6^{(1)} [\dot{v}_4 + K(v_4 + \tau\dot{v}_3)] = Q^{(1)},\\
&\dot{n}_1^{(1)} e^{\sigma\tau} w_1 - \dot{n}_2^{(1)} e^{-\sigma\tau} w_2 + \dot{n}_3^{(1)} e^{\bar{\sigma}\tau} \bar{w}_1 - \dot{n}_4^{(1)} e^{-\bar{\sigma}\tau} \bar{w}_2 + \dot{n}_5^{(1)} w_3 + \dot{n}_6^{(1)} (w_4 + K\tau w_3) = 0,\\
&\dot{n}_1^{(1)} e^{\sigma\tau} (\sigma w_1 + \dot{w}_1) + \dot{n}_2^{(1)} e^{-\sigma\tau} (\sigma w_2 - \dot{w}_2) + \dot{n}_3^{(1)} e^{\bar{\sigma}\tau} (\bar{\sigma}\bar{w}_1 + \dot{\bar{w}}_1) + \dot{n}_4^{(1)} e^{-\bar{\sigma}\tau} (\bar{\sigma}\bar{w}_2 - \dot{\bar{w}}_2)\\
&\qquad + \dot{n}_5^{(1)} \dot{w}_3 + \dot{n}_6^{(1)} [\dot{w}_4 + K(w_4 + \tau\dot{w}_3)] = R^{(1)}.
\end{aligned}\right\} \quad \ldots(67)$$

The determinant of the coefficients of $\dot{n}_1^{(1)}, \ldots, \dot{n}_6^{(1)}$ in the above equations is $\Delta$, the same as in (60), and is different from zero for $\epsilon$ not zero but sufficiently small numerically, and for all $\mu$ such that $\mu_0 \leqq \mu \leqq \frac{1}{2}$. Hence equations (67) can be solved for $\dot{n}_1^{(1)}, \ldots, \dot{n}_6^{(1)}$, the solutions being

$$\dot{n}_j^{(1)} = \frac{\Delta_j^{(1)}}{\Delta}, \; (j = 1, 2, \ldots, 6), \quad \ldots\ldots(68)$$

where $\Delta_j^{(1)}$ is the determinant formed by replacing the elements of the $j$th column of $\Delta$ by 0, $P^{(1)}$, 0, $Q^{(1)}$, 0, and $R^{(1)}$, respectively. Since $P^{(1)}$, $Q^{(1)}$, and $R^{(1)}$ do not contain terms in $e^{\pm\sigma\tau}$ or $e^{\pm\bar{\sigma}\tau}$, the integrations of (68) for $\dot{n}_1^{(1)}, \ldots, \dot{n}_4^{(1)}$ will yield no terms in $\tau$ explicitly. Terms in $\tau$ will occur, however, in the integrations for $n_5^{(1)}$ and $n_6^{(1)}$, but when the values for $n_5^{(1)}$ and $n_6^{(1)}$ are substituted in the complementary functions the terms in $\tau$ which arise from the particular integrals cancel off. By substituting the values of $n_1^{(1)}, \ldots, n_6^{(1)}$ in the complementary function, we obtain the complete solutions. They are thus found to be

$$\left.\begin{aligned}
p_1 &= N_1^{(1)} e^{\sigma\tau} u_1 + N_2^{(1)} e^{-\sigma\tau} u_2 + N_3^{(1)} e^{\bar{\sigma}\tau} \bar{u}_1 + N_4^{(1)} e^{-\bar{\sigma}\tau} \bar{u}_2 + N_5^{(1)} u_3\\
&\quad + N_6^{(1)} (u_4 + K\tau u_3) + \epsilon a^2 [e^{-2\sigma\tau} u_{20}^{(1)} + e^{-(\sigma+\bar{\sigma})\tau} u_{11}^{(1)} + e^{-2\bar{\sigma}\tau} u_{02}^{(1)}],\\
q_1 &= N_1^{(1)} e^{\sigma\tau} v_1 - N_2^{(1)} e^{-\sigma\tau} v_2 + N_3^{(1)} e^{\bar{\sigma}\tau} \bar{v}_1 - N_4^{(1)} e^{-\bar{\sigma}\tau} \bar{v}_2 + N_5^{(1)} v_3\\
&\quad + N_6^{(1)} (v_4 + K\tau v_3) + \epsilon a^2 [e^{-2\sigma\tau} v_{20}^{(1)} + e^{-(\sigma+\bar{\sigma})\tau} v_{11}^{(1)} + e^{-2\bar{\sigma}\tau} v_{02}^{(1)}],\\
r_1 &= N_1^{(1)} e^{\sigma\tau} w_1 - N_2^{(1)} e^{-\sigma\tau} w_2 + N_3^{(1)} e^{\bar{\sigma}\tau} \bar{w}_1 - N_4^{(1)} e^{-\bar{\sigma}\tau} \bar{w}_2 + N_5^{(1)} w_3\\
&\quad + N_6^{(1)} (w_4 + K\tau w_3) + \epsilon a^2 [e^{-2\sigma\tau} w_{20}^{(1)} + e^{-(\sigma+\sigma)\tau} w_{11}^{(1)} + e^{-2\bar{\sigma}\tau} w_{02}^{(1)}],
\end{aligned}\right\} \quad \ldots\ldots(69)$$

where $N_1^{(1)}, \ldots, N_6^{(1)}$ are the constants of integration: and $u^{(1)}$, $v^{(1)}$, and $w^{(1)}$ with the various subscripts are power series of the same form, respectively, as $U^{(1)}$, $V^{(1)}$, and $W^{(1)}$ with the same subscripts. When condition $(C_1)$ is imposed on (69), then

$$N_1^{(1)} = N_3^{(1)} = N_5^{(1)} = N_6^{(1)} = 0.$$

The remaining constants $N_2^{(1)}$ and $N_4^{(1)}$ can be uniquely determined by the conditions $(C_2)$. As at the previous step, this determination of $N_2^{(1)}$ and $N_4^{(1)}$ will yield values which are conjugate complex. The desired solutions which satisfy both conditions $(C_1)$ and $(C_2)$ are therefore

$$\left.\begin{aligned} p_1 &= \epsilon a^2 \left[e^{-\sigma\tau} u_{10}^{(1)} + e^{-\bar{\sigma}\tau} u_{01}^{(1)} + e^{-2\sigma\tau} u_{20}^{(1)} + e^{-(\sigma+\bar{\sigma})\tau} u_{11}^{(1)} + e^{-2\bar{\sigma}\tau} u_{02}^{(1)}\right], \\ q_1 &= \epsilon a^2 \left[e^{-\sigma\tau} v_{10}^{(1)} + e^{-\bar{\sigma}\tau} v_{01}^{(1)} + e^{-2\sigma\tau} v_{20}^{(1)} + e^{-(\sigma+\bar{\sigma})\tau} v_{11}^{(1)} + e^{-2\bar{\sigma}\tau} v_{02}^{(1)}\right], \\ r_1 &= \epsilon a^2 \left[e^{-\sigma\tau} w_{10}^{(1)} + e^{-\bar{\sigma}\tau} w_{01}^{(1)} + e^{-2\sigma\tau} w_{20}^{(1)} + e^{-(\sigma+\bar{\sigma})\tau} w_{11}^{(1)} + e^{-2\bar{\sigma}\tau} w_{02}^{(1)}\right], \end{aligned}\right\} \quad \text{.........(70)}$$

where $u_{10}^{(1)}$, $v_{10}^{(1)}$, and $w_{10}^{(1)}$ are of the same form as $u_1$, $v_1$, and $w_1$ respectively; and

$$u_{01}^{(1)} = \bar{u}_{10}^{(1)}, \quad v_{01}^{(1)} = \bar{v}_{10}^{(1)}, \quad w_{01}^{(1)} = \bar{w}_{10}^{(1)}.$$

The remaining steps of the integration can be carried on in precisely the same way. By an induction to the general term we shall show that the integration can be carried on as far as is desired.

Let us suppose that $p_\nu$, $q_\nu$, and $r_\nu$, $(\nu = 1, \ldots, m-1)$, have all been determined, and that

$$\left.\begin{aligned} p_\nu &= \epsilon^\nu a^{\nu+1} \sum_{k=1}^{\nu+1} \sum_{j=0}^{k} e^{-[(k-j)\sigma + j\bar{\sigma}]\tau} u_{k-j,j}^{(\nu)}, \\ q_\nu &= \epsilon^\nu a^{\nu+1} \sum_{k=1}^{\nu+1} \sum_{j=0}^{k} e^{-[(k-j)\sigma + j\bar{\sigma}]\tau} v_{k-j,j}^{(\nu)}, \\ r_\nu &= \epsilon^\nu a^{\nu+1} \sum_{k=1}^{\nu+1} \sum_{j=0}^{k} e^{-[(k-j)\sigma + j\bar{\sigma}]\tau} w_{k-j,j}^{(\nu)}, \end{aligned}\right\} \quad \text{.........................(71)}$$

where $u_{k-j,j}^{(\nu)}$, $v_{k-j,j}^{(\nu)}$, and $w_{k-j,j}^{(\nu)}$ are power series of the same form respectively as $u_1$, $v_1$, and $w_1$ if $j \neq \frac{1}{2}k$. If $k$ is even and $j = \frac{1}{2}k$, then the functions in (71) are of the same form as $u_1$, $v_1$, and $w_1$, respectively, except that the coefficients of the sines and cosines are all real instead of complex. Further,

$$u_{k_1 k_2}^{(\nu)} = \bar{u}_{k_2 k_1}^{(\nu)}, \quad v_{k_1 k_2}^{(\nu)} = \bar{v}_{k_2 k_1}^{(\nu)}, \quad w_{k_1 k_2}^{(\nu)} = \bar{w}_{k_2 k_1}^{(\nu)}, \quad k_1 + k_2 = k.$$

We propose to show that $p_\nu$, $q_\nu$, and $r_\nu$ have the same form as (71) when $\nu = m$. Consider the set of differential equations obtained by equating the coefficients of $\gamma^m$ in (43′), that is in (43) after (62) and (63) have been substituted and $\gamma$ has been divided out. Except for the subscript $m$, these differential equations have the same left members as the equations of variation. Let the right members be denoted by $P^{(m)}$, $Q^{(m)}$, and $R^{(m)}$ respectively. Then

$$P^{(m)} = \epsilon^m a^{m+1} \sum_{k=2}^{m+1} \sum_{j=0}^{k} e^{-[(k-j)\sigma + j\bar{\sigma}]\tau} U_{k-j,j}^{(m)},$$

$$Q^{(m)} = \epsilon^m a^{m+1} \sum_{k=2}^{m+1} \sum_{j=0}^{k} e^{-[(k-j)\sigma + j\bar{\sigma}]\tau} V_{k-j,j}^{(m)},$$

$$R^{(m)} = \epsilon^m a^{m+1} \sum_{k=2}^{m+1} \sum_{j=0}^{k} e^{-[(k-j)\sigma - j\bar{\sigma}]\tau} W_{k-j,j}^{(m)},$$

where $U_{k-j,j}^{(m)}$, $V_{k-j,j}^{(m)}$, and $W_{k-j,j}^{(m)}$ are similar to $u_{k-j,j}^{(\nu)}$, $v_{k-j,j}^{(\nu)}$, and $w_{k-j,j}^{(\nu)}$ respectively.

The complementary functions of the differential equations in $p_m$, $q_m$, and $r_m$ are the same as (61). If we denote the arbitrary constants by $n_1^{(m)}, \ldots, n_6^{(m)}$ and employ the method of the variation of parameters as in the preceding steps, we obtain equations the same as (67) except for the superscript $m$. Since the right members $P^{(m)}$, $Q^{(m)}$, and $R^{(m)}$ do not contain terms in $e^{\pm\sigma\tau}$ or $e^{\pm\bar{\sigma}\tau}$, the integrations of the equations analogous to (68) for $n_1^{(m)}, \ldots, n_4^{(m)}$ will not contain terms in $\tau$ explicitly. Proceeding as in the determination of $p_1$, $q_1$, and $r_1$, we find

$$p_m = N_1^{(m)} e^{\sigma\tau} u_1 + N_2^{(m)} e^{-\sigma\tau} u_2 + N_3^{(m)} e^{\bar{\sigma}\tau} \bar{u}_1 + N_4^{(m)} e^{-\bar{\sigma}\tau} \bar{u}_2 + N_5^{(m)} u_3$$
$$+ N_6^{(m)} (u_4 + K\tau u_3) + \epsilon^m a^{m+1} \sum_{k=2}^{m+1} \sum_{j=0}^{k} e^{-[(k-j)\sigma + j\bar{\sigma}]\tau} u_{k-j,j}^{(m)},$$

$$q_m = N_1^{(m)} e^{\sigma\tau} v_1 - N_2^{(m)} e^{-\sigma\tau} v_2 + N_3^{(m)} e^{\bar{\sigma}\tau} \bar{v}_1 - N_4^{(m)} e^{-\bar{\sigma}\tau} \bar{v}_2 + N_5^{(m)} v_3$$
$$+ N_6^{(m)} (v_4 + K\tau v_3) + \epsilon^m a^{m+1} \sum_{k=2}^{m+1} \sum_{j=0}^{k} e^{-[(k-j)\sigma + j\bar{\sigma}]\tau} v_{k-j,j}^{(m)},$$

$$r_m = N_1^{(m)} e^{\sigma\tau} w_1 - N_2^{(m)} e^{-\sigma\tau} w_2 + N_3^{(m)} e^{\bar{\sigma}\tau} \bar{w}_1 - N_4^{(m)} e^{-\bar{\sigma}\tau} \bar{w}_2 + N_5^{(m)} w_3$$
$$+ N_6^{(m)} (w_4 + K\tau w_3) + \epsilon^m a^{m+1} \sum_{k=2}^{m+1} \sum_{j=0}^{k} e^{-[(k-j)\sigma + j\bar{\sigma}]\tau} w_{k-j,j}^{(m)},$$

where $N_1^{(m)}, \ldots, N_6^{(m)}$ are the constants of integration, and the functions $u_{k-j,j}^{(m)}$, $v_{k-j,j}^{(m)}$, and $w_{k-j,j}^{(m)}$ are similar to the corresponding functions in (71) with the superscripts $\nu$. When conditions ($C_1$) and ($C_2$) are imposed on the above solutions, it is found that

$$N_1^{(m)} = N_3^{(m)} = N_5^{(m)} = N_6^{(m)} = 0,$$

and that $N_2^{(m)}$ and $N_4^{(m)}$ are conjugate complex numbers which carry the factor $\epsilon^m a^{m+1}$. With this determination of the constants of integration, the solutions for $p_m$, $q_m$, and $r_m$ have the same form as (71) if $\nu = m$. This completes the induction to the general term.

Returning to the variables $p$, $q$, and $r$ by means of (63) and (62), we find that the asymptotic solutions of (43) which approach zero as $\tau$ approaches $+\infty$ are

$$\left.\begin{aligned} p &= p_0\gamma + p_1\gamma^2 + \ldots + p_n\gamma^{n+1} + \ldots, \\ q &= q_0\gamma + q_1\gamma^2 + \ldots + q_n\gamma^{n+1} + \ldots, \\ r &= r_0\gamma + r_1\gamma^2 + \ldots + r_n\gamma^{n+1} + \ldots, \end{aligned}\right\} \quad \ldots\ldots\ldots\ldots\ldots(72)$$

where the various $p_\nu$, $q_\nu$, and $r_\nu$ ($\nu = 0, 1, \ldots \infty$) are defined in (71). If we change the superscripts in (71), so as to correspond with the powers of $\gamma$ in (72), these solutions may be written

$$\left.\begin{aligned} p &= \sum_{\nu=0}^{\infty} \sum_{k=1}^{\nu+1} \sum_{j=0}^{k} e^{-[(k-j)\sigma + j\bar{\sigma}]\tau} u_{k-j,j}^{(\nu+1)} \epsilon^\nu (a\gamma)^{\nu+1}, \\ q &= \sum_{\nu=0}^{\infty} \sum_{k=1}^{\nu+1} \sum_{j=0}^{k} e^{-[(k-j)\sigma + j\bar{\sigma}]\tau} v_{k-j,j}^{(\nu+1)} \epsilon^\nu (a\gamma)^{\nu+1}, \\ r &= \sum_{\nu=0}^{\infty} \sum_{k=1}^{\nu+1} \sum_{j=0}^{k} e^{-[(k-j)\sigma + j\bar{\sigma}]\tau} w_{k-j,j}^{(\nu+1)} \epsilon^\nu (a\gamma)^{\nu+1}. \end{aligned}\right\} \quad \ldots\ldots\ldots\ldots\ldots(73)$$

Since $a$ and $\gamma$ occur only in products, as indicated in (73), and both are arbitrary, we may put either equal to 1 without loss of generality. We shall therefore consider $a$ to be 1 in (73).

The coefficients of the sines and cosines in the various functions $u_{k-j,j}^{(\nu+1)}$, $v_{k-j,j}^{(\nu+1)}$, and $w_{k-j,j}^{(\nu+1)}$ are complex except when $k$ is even and $j = \frac{1}{2}k$. Since we desire a *real* asymptotic orbit, it is necessary to show that the solutions (73) are real.

With each term

$$e^{-(k_1\sigma+k_2\bar{\sigma})\tau}u^{(\nu+1)}_{k_1k_2} \quad \text{.........................................(74)}$$

of (73), there may be associated the term

$$e^{-(k_2\sigma+k_1\bar{\sigma})\tau}u^{(\nu+1)}_{k_2k_1}, \quad \text{.........................................(75)}$$

unless $k_1=k_2$. Suppose first $k_1\neq k_2$. Then the coefficients of the sines and cosines of the various powers of $\epsilon$ in $u^{(\nu+1)}_{k_1k_2}$ are the conjugates of the corresponding coefficients in $u^{(\nu+1)}_{k_2k_1}$. When the two expressions (74) and (75) are added, we obtain terms of the type

$$e^{-(k_1\sigma+k_2\bar{\sigma})\tau}[A\cos j\tau+B\sin j\tau]+e^{-(k_2\sigma+k_1\bar{\sigma})\tau}[\bar{A}\cos j\tau+\bar{B}\sin j\tau]. \quad \text{...........(76)}$$

Since $\sigma=\alpha+i\beta$,

$$\left.\begin{aligned} e^{-(k_1\sigma+k_2\bar{\sigma})\tau}&=e^{-(k_1+k_2)\alpha\tau}[\cos(k_1-k_2)\beta\tau-i\sin(k_1-k_2)\beta\tau],\\ e^{-(k_2\sigma+k_1\bar{\sigma})\tau}&=e^{-(k_1+k_2)\alpha\tau}[\cos(k_1-k_2)\beta\tau+i\sin(k_1-k_2)\beta\tau].\end{aligned}\right\} \quad \text{...............(77)}$$

Now let $A=a_1+ib_1$, $B=a_2+ib_2$. Then by virtue of the relations (77), the expression (76) becomes

$$\begin{aligned} e^{-(k_1+k_2)\alpha\tau}[(a_1-b_2)\cos\{(k_1-k_2)\beta+j\}\tau+(a_1+b_2)\cos\{(k_1-k_2)\beta+j\}\tau\\ +(a_2+b_1)\sin\{(k_1-k_2)\beta+j\}\tau+(b_1-a_2)\sin\{(k_1-k_2)\beta+j\}\tau],\end{aligned}$$

which is real since all the constants are real.

If $k_1=k_2=k$, then the coefficients in $u^{(\nu+1)}_{kk}$ are all real. In this case (74) consists of terms of the type

$$e^{-2k\alpha\tau}[a_1\cos j\tau+b_1\sin j\tau],$$

which are real. Therefore the solutions (73) are real.

In order to express the solutions (73) in a form in which the imaginaries will not appear, we may put for $p_\nu$, $q_\nu$, and $r_\nu$ of (71)

$$\left.\begin{aligned} p_\nu&=e^\nu a^{\nu+1}\sum_{l=1}^{\nu+1}e^{-l\alpha\tau}T^{(\nu+1)}_{1l},\\ q_\nu&=e^\nu a^{\nu+1}\sum_{l=1}^{\nu+1}e^{-l\alpha\tau}T^{(\nu+1)}_{2l},\\ r_\nu&=e^\nu a^{\nu+1}\sum_{l=1}^{\nu+1}e^{-l\alpha\tau}T^{(\nu+1)}_{3l},\end{aligned}\right\} \quad \text{....................................(78)}$$

where if $l$ is even, $l=2k$ say, we have, for $j_1=1, 2, 3$,

$$\begin{aligned} T^{(\nu+1)}_{j_1l}=\sum_{j=0}^{\infty}\sum_{k_1=0}^{k}[C^{(\nu+1,2j)}_{j_1,+2k_1}\cos 2(j+k_1\beta)\tau+C^{(\nu+1,2j)}_{j_1,-2k_1}\cos 2(j-k_1\beta)\tau\\ +S^{(\nu+1,2j)}_{j_1,+2k_1}\sin 2(j+k_1\beta)\tau+S^{(\nu+1,2j)}_{j_1,-2k_1}\sin 2(j-k_1\beta)\tau]\epsilon^{2j}.\end{aligned}$$

If $l$ is odd, $l=2k+1$, the values of $T^{(\nu+1)}_{j_1l}$, $(j_1=1, 2, 3)$, are obtained from the preceding by replacing $2k_1$ with $2k_1+1$. The coefficients of the cosines and sines in the above equations are all real.

Returning to the original coordinates $\xi$, $\eta$, $\zeta$ through the substitutions (42), (3), and (2), we obtain

$$\left.\begin{aligned} \xi&=+\tfrac{1}{2}-\mu+\epsilon x_0(+\sqrt{3},\tau)+\epsilon p(+\sqrt{3},\tau),\\ \eta&=+\tfrac{1}{2}\sqrt{3}+\epsilon y_0(+\sqrt{3},\tau)+\epsilon q(+\sqrt{3},\tau),\\ \zeta&=\qquad\epsilon z_0(+\sqrt{3},\tau)+\epsilon r(+\sqrt{3},\tau)\end{aligned}\right\} \quad \text{........................(79, I)}$$

as the parametric representation of the orbit which approaches the periodic orbit about the point I as $\tau$ approaches $+\infty$. The corresponding orbit near the equilibrium point II is obtained by changing the sign of $\sqrt{3}$ wherever it occurs in (79, I). This orbit is

$$\left.\begin{aligned}\xi &= \tfrac{1}{2}-\mu+\epsilon x_0(-\sqrt{3},\tau)+\epsilon p(-\sqrt{3},\tau),\\ \eta &= -\tfrac{1}{2}\sqrt{3}+\epsilon y_0(-\sqrt{3},\tau)+\epsilon q(-\sqrt{3},\tau),\\ \zeta &= \epsilon z_0(-\sqrt{3},\tau)+\epsilon r(-\sqrt{3},\tau).\end{aligned}\right\} \quad \text{(79, II)}$$

(B) *Solutions in* $e^{a\tau}$.

Let us next consider the orbits which approach the equilibrium points as $\tau$ approaches $-\infty$. The construction of such orbits would be similar to the preceding construction. The condition ($C_1$) would be altered so that each term of the solutions of (43) would contain the factor $e^{\sigma\tau}$ or $e^{\bar{\sigma}\tau}$, as these exponentials have their real parts positive. The asymptotic solutions corresponding to (78) would then contain powers of $e^{a\tau}$ instead of $e^{-a\tau}$.

It is not necessary to consider the construction of these orbits in detail, however, as they may be obtained from the preceding by changing the signs of $\sqrt{3}$, $\tau$, $q$, and $r$ in (73). As was shown in § 8, the differential equations (43) remain unchanged by changing the signs of $\sqrt{3}$, $\tau$, $q$, and $r$, and consequently the same changes in the solutions will still leave them solutions. If we denote (73) by $p(+\sqrt{3},\tau)$, $q(+\sqrt{3},\tau)$, and $r(+\sqrt{3},\tau)$, then $p(-\sqrt{3},-\tau)$, $-q(-\sqrt{3},-\tau)$, and $-r(-\sqrt{3},-\tau)$ will also be solutions of (43). The asymptotic orbit which approaches the point I as $\tau$ approaches $-\infty$ is therefore

$$\left.\begin{aligned}\xi &= \tfrac{1}{2}-\mu+\epsilon x_0(+\sqrt{3},\tau)+\epsilon p(-\sqrt{3},-\tau),\\ \eta &= +\tfrac{1}{2}\sqrt{3}+\epsilon y_0(+\sqrt{3},\tau)-\epsilon q(-\sqrt{3},-\tau),\\ \zeta &= \epsilon z_0(+\sqrt{3},\tau)-\epsilon r(-\sqrt{3},-\tau).\end{aligned}\right\} \quad \text{(80, I)}$$

The corresponding orbit for the point II is obtained from the above by changing the sign of $\sqrt{3}$. This orbit is

$$\left.\begin{aligned}\xi &= \tfrac{1}{2}-\mu+\epsilon x_0(-\sqrt{3},\tau)+\epsilon p(+\sqrt{3},-\tau),\\ \eta &= -\tfrac{1}{2}\sqrt{3}+\epsilon y_0(-\sqrt{3},\tau)-\epsilon q(+\sqrt{3},-\tau),\\ \zeta &= \epsilon z_0(-\sqrt{3},\tau)-\epsilon r(+\sqrt{3},-\tau).\end{aligned}\right\} \quad \text{(80, II)}$$

## 10. THE UNDETERMINED CONSTANTS.

The asymptotic solutions (79, I), (79, II), (80, I), and (80, II) contain three arbitrary parameters, viz. $t_0$, $\epsilon$, and $\gamma$. The constant $t_0$ represents the initial time and may be put equal to zero without loss of generality. The parameter $\epsilon$ enters in the construction of the periodic orbits and is the scale factor for these orbits. From the way in which the initial values for the periodic solutions were chosen, the parameter $\epsilon$ is found to be proportional to the initial projection of the infinitesimal body from the plane of motion of the finite bodies, the initial projection being $\epsilon/(1+\delta)$. The parameter $\gamma$ is the scale factor for the asymptotic orbits. From the choice of the initial conditions ($C_2$) it follows that $\gamma$ represents the initial displacement of the infinitesimal body from the periodic orbit on a line parallel to the $\xi$-axis and in the $\xi\eta$-plane. Since this constant enters the asymptotic solutions both to even and odd degrees, then the orbit obtained by taking a positive value of $\gamma$ is not only differently orientated but is geometrically distinct from the orbit obtained by taking the same numerical value of $\gamma$ but the opposite sign.

From an examination of the derivative $\frac{dy}{dx}$ it is found, as in Part I, that the directions in which the asymptotic orbits approach the periodic orbits are indeterminate.

### 10*a*. Numerical Example.

To illustrate the nature of the periodic orbits and the corresponding asymptotic orbits, we have assigned to $\mu$ the value 0·1, as in Part I, and to $\epsilon$ the value 0·5. With these values of $\mu$ and $\epsilon$ the periodic solutions (39), up to and including the terms in $\epsilon^3$, are

$$\left.\begin{aligned}\bar{\bar{x}} &= \epsilon x_0 = 0{\cdot}0238 \cos 2\tau + 0{\cdot}0514 \sin 2\tau,\\ \bar{\bar{y}} &= \epsilon y_0 = -0{\cdot}0723 + 0{\cdot}0550 \cos 2\tau - 0{\cdot}0238 \sin 2\tau,\\ \bar{\bar{z}} &= \epsilon z_0 = 0{\cdot}4985 \sin \tau + 0{\cdot}0005 \sin 3\tau.\end{aligned}\right\} \qquad (81)$$

Numerical values of these coordinates for various values of $\tau$ are to be found in Table III.

Table III.

$\mu = 0{\cdot}1$ $\qquad$ $\epsilon = 0{\cdot}5$

| $\tau$ | $\epsilon x_0$ | $\epsilon y_0$ | $\epsilon z_0$ |
|---|---|---|---|
| 0 | + ·0238 | − ·0173 | 0 |
| ·1 | + ·0335 | − ·0231 | + ·0499 |
| ·2 | + ·0419 | − ·0310 | + ·0993 |
| ·3 | + ·0486 | − ·0403 | + ·1472 |
| ·4 | + ·0535 | − ·0511 | + ·1941 |
| ·5 | + ·0561 | − ·0626 | + ·2390 |
| 1 | + ·0368 | − ·1168 | + ·4197 |
| 1·5 | − ·0164 | − ·1301 | + ·4980 |
| 2 | − ·0545 | − ·0903 | + ·4536 |
| 2·5 | − ·0425 | − ·0339 | + ·2984 |
| 3 | + ·0085 | − ·0129 | + ·0704 |
| 3·5 | + ·0517 | − ·0464 | − ·1751 |
| 4 | + ·0473 | − ·1038 | − ·3777 |
| 4·5 | − ·0005 | − ·1322 | − ·4875 |
| 5 | − ·0480 | − ·1055 | − ·4785 |
| 5·5 | + ·0235 | − ·0172 | − ·3523 |
| 6 | − ·0075 | − ·0131 | − ·1392 |
| 6·5 | + ·0432 | − ·0234 | + ·1073 |
| 7 | + ·0542 | − ·0884 | + ·3278 |
| 7·5 | + ·0153 | − ·1296 | + ·4681 |
| 8 | − ·0376 | − ·1181 | + ·4935 |
| 8·5 | − ·0559 | − ·0645 | + ·3987 |
| 9 | − ·0229 | − ·0181 | + ·2056 |
| 9·5 | + ·0312 | − ·0215 | − ·0374 |
| 10 | + ·0566 | − ·0716 | − ·2715 |

The projections on the coordinate planes of the periodic orbit (81) near the equilibrium point I are given in the heavy lines of Figs. 3, 4, and 5*. The orbit consists of two loops, one above and the other below the $xy$-plane, with the double point in the fourth quadrant of the

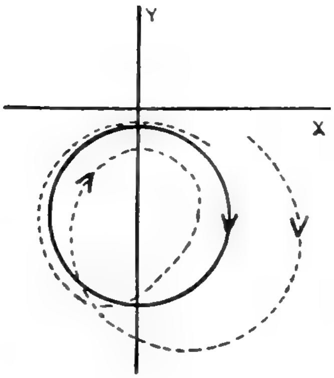



FIG. 3.

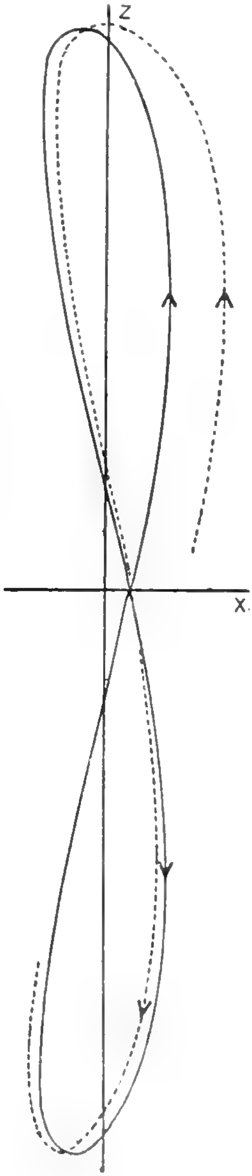



FIG. 4.

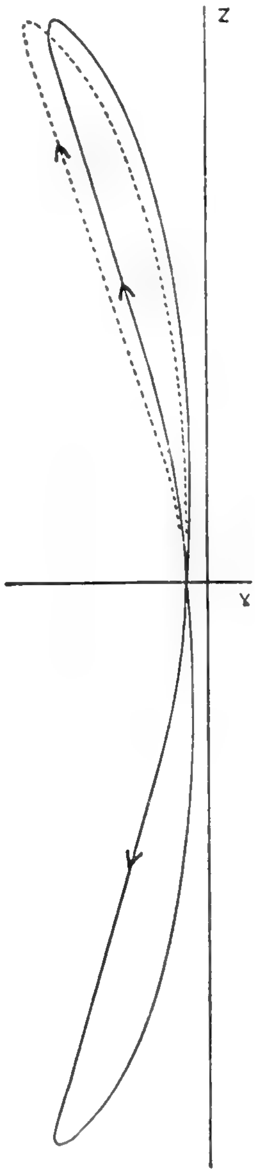



FIG. 5.

Projections on the coordinate planes of the Periodic and Asymptotic Orbits near the Equilibrium Point I.

$xy$-plane. The projection on this plane is approximately a circle with centre at $(0, -0{\cdot}0723)$ and radius $0{\cdot}06$ (see Fig. 3).

To obtain an orbit that is asymptotic to the above periodic orbit, we consider particular values of $p$, $q$, and $r$ which are added to $x_0$, $y_0$, and $z_0$, respectively, in (42). After $\mu$ and $\epsilon$ are

* See also Figs. 5, 6 and 7 of "Oscillating Satellite."

fixed, the only undetermined constant in $p$, $q$, and $r$ is the scale factor $\gamma$ which was introduced in equations (62). If $\gamma$ is put equal to 0·1, and only the first terms of (72) are taken, then the solutions for $\epsilon p$, $\epsilon q$, and $\epsilon r$ in (79, I) are

$$\epsilon p = e^{-\alpha\tau}(0{\cdot}05 \cos\beta\tau + 0{\cdot}0036 \sin\beta\tau),$$
$$\epsilon q = e^{-\alpha\tau}(-0{\cdot}0525 \sin\beta\tau),$$
$$\epsilon r = e^{-\alpha\tau}\{\sin\tau(0{\cdot}0385\cos\beta\tau - 0{\cdot}0080\sin\beta\tau) + \cos\tau(0{\cdot}0205\cos\beta\tau - 0{\cdot}0865\sin\beta\tau)\}.$$

These solutions represent the $x$-, $y$- and $z$-components, respectively, of the amount which the asymptotic orbit deviates from the periodic orbit. Numerical values of these displacements are given in Table IV.

TABLE IV.

$\mu = 0{\cdot}1$ $\epsilon = 0{\cdot}5$ $\gamma = 0{\cdot}1$

| $\tau$ | $\epsilon p$ | $\epsilon q$ | $\epsilon r$ |
|---|---|---|---|
| 0 | + ·0500 | 0 | + ·0205 |
| ·1 | + ·0507 | − ·0040 | + ·0166 |
| ·2 | + ·0512 | − ·0075 | + ·0126 |
| ·3 | + ·0511 | − ·0110 | + ·0088 |
| ·4 | + ·0502 | − ·0140 | + ·0053 |
| ·5 | + ·0499 | − ·0170 | + ·0022 |
| 1 | + ·0419 | − ·0255 | − ·0063 |
| 1·5 | + ·0296 | − ·0280 | + ·0008 |
| 2 | + ·0165 | − ·0250 | + ·0140 |
| 2·5 | + ·0048 | − ·0185 | + ·0204 |
| 3 | − ·0041 | − ·0115 | + ·0222 |
| 3·5 | − ·0095 | − ·0050 | + ·0159 |
| 4 | − ·0117 | + ·0005 | + ·0087 |
| 4·5 | − ·0115 | + ·0045 | + ·0049 |
| 5 | − ·0090 | + ·0060 | + ·0053 |
| 5·5 | − ·0065 | + ·0065 | + ·0073 |
| 6 | − ·0030 | + ·0055 | + ·0087 |
| 6·5 | − ·0005 | + ·0035 | + ·0079 |
| 7 | + ·0010 | + ·0025 | + ·0055 |
| 7·5 | + ·0025 | + ·0010 | + ·0032 |
| 8 | + ·0025 | − ·0005 | + ·0018 |
| 8·5 | + ·0025 | − ·0010 | + ·0016 |
| 9 | + ·0021 | − ·0015 | + ·0020 |
| 9·5 | + ·0015 | − ·0015 | + ·0022 |
| 10 | + ·0006 | − ·0001 | + ·0019 |

The dotted lines in Figs. 3, 4, and 5 represent the projections on the coordinate planes of the asymptotic orbit. In Figs. 3 and 4 the asymptotic orbits are drawn as $\tau$ varies from 0 to $2\pi$, approximately, that is for one period of the closed orbits. In Fig. 5 the deviation of the asymptotic orbit from the closed orbit is very small as $\tau$ varies from $\pi$ to $2\pi$, and is not represented in the drawing. At each succeeding period the asymptotic orbit lies between the corresponding branches of the periodic and asymptotic orbits for the same phase of the preceding period.

In conclusion the author desires to express his thanks to his colleague K. P. Johnston, B.A., B.Sc., for making the drawings which appear in this paper.

## XVI. *Terrestrial Magnetic Variations and their connection with Solar Emissions which are Absorbed in the Earth's Outer Atmosphere.*

By S. Chapman, M.A., D.Sc., F.R.S., Trinity College.

[*Read* 17 February 1919.]

### INTRODUCTION.

§1. The aim of this paper is to abstract and interpret various features of the non-secular changes of the earth's magnetism which seem to be of special significance in connection with atmospheric and solar physics.

The non-secular magnetic variations are everywhere small compared with the total intensity of the earth's field. Magnetographs installed at a number of widely distributed observatories provide a continuous record of the magnetic changes, resolved in three directions. At any one station the traces show that there exists a well-marked diurnal variation of the force vector, but that the course of this variation is ordinarily complicated by superposed irregular perturbations. At times the latter are unusually intense or frequent, at others they are almost absent. It is customary to classify each day according to the amount of this *disturbance* (as it is called) either by the terms quiet, ordinary, and disturbed, or by corresponding 'character' figures 0, 1, 2. No precise canons of classification have so far been arrived at, and a merely three-fold subdivision must obviously be very rough. Also, since the ordinary or average amount of disturbance varies with locality and season, a character figure 0 (say) will correspond to different absolute amounts of perturbation at different stations. But it is found that on the whole a quiet day* at one station is quiet also at most others, *i.e.* that magnetic conditions all over the earth are generally similar as regards the degree of disturbance existent, relative to the normal amount characteristic of each region. By international agreement a mean character figure is derived for every Greenwich day, from the figures independently assigned at each of a large number of cooperating observatories. The five days in each month which have the smallest mean character figures are termed the 'international quiet days,' and it is customary for observatories to publish monthly mean hourly inequalities of the three elements of magnetic force from all (or all but highly disturbed) days of each month, and in addition from the five quiet days only†.

These hourly inequalities indicate the amplitude and type of the diurnal magnetic variations. The influence of disturbance on the latter can be to some extent inferred by a comparison of the inequalities derived from all and from quiet days. The quiet days themselves, however, cannot be regarded as wholly free from the effects of some small present disturbance, or from the after-effects of past disturbance. But when the nature of disturbance effects has become clear, allowance can be made for the residual amount on the five 'quiet' days, and a conception formed

* Reckoned from one Greenwich midnight to the next.

† Before the international arrangement came into being, the Astronomer Royal chose five quiet days monthly from the Greenwich magnetic records, and the monthly mean diurnal variations from these days have been published at Greenwich since 1889.

of the phenomena which would characterize ideally quiet conditions. These phenomena will be termed the *regular diurnal variations*, and the magnetic variations deduced by abstracting these regular variations from the changes actually observed will be termed the *disturbance variations*. The latter are not wholly irregular, and in particular they include certain additional diurnal variations, superposed during disturbed periods upon those characteristic of the ideally quiet state, which will be called *disturbance diurnal variations*. These are true diurnal variations, and are to be distinguished from such quite different things as, *e.g.* the diurnal variation of *frequency* of disturbance.

For a first approximation, the regular diurnal variations can generally be identified with those derived from the five quietest days per month. As regards disturbance, its effects are most clearly marked during the violent outbreaks to which Humboldt gave the name *magnetic storm*. The discussion first deals with disturbance in this form, proceeding thence to consider more ordinary disturbance, and afterwards describing the phenomena of quiet days; the relation of these three classes of variation to locality upon the earth and to the geocentric coordinates of the sun is described in §§2–8, and afterwards, in §§9–13, their relation to the physical condition of the sun. A partial interpretation of the phenomena is attempted in §§14–23.

## MAGNETIC DISTURBANCE.

§2. In a recent paper* I have discussed the changes which occur during magnetic storms. Besides irregular fluctuations, most numerous and intense in high latitudes, there are more or less regular variations, which appear clearly on averaging the changes observed during a number of storms. These regular effects depend on terrestrial latitude and on local time, *i.e.* upon longitude reckoned from the meridian containing the sun. They can be analyzed into changes depending on latitude only (*i.e.* the mean change along any circle of latitude) and the residual changes which also depend on local time. The latter appear as additions to the regular diurnal variations; they are, in the present terminology, the disturbance diurnal variations. The other regular changes, common to all longitudes, are of a simple definite type, corresponding mainly to a slight demagnetization of the horizontal magnetic field, during the first few hours of the storm, with subsequent slow recovery; they therefore depend on 'storm time,' reckoned from the commencement of the storm, which is nearly simultaneous all over the earth. The disturbance diurnal variations wax and wane in amplitude in unison with the storm-time variations; they merely represent, indeed, an inequality in the intensity of the latter in different regions, the P.M. hemisphere being more affected than the A.M. hemisphere†. This inequality is of simple type, there being one maximum and one minimum in the regular storm effects along any circle of latitude; these occur approximately at $18^h$ and $6^h$ local time, respectively. The disturbance diurnal variations are consequently almost purely diurnal sine waves, and their phases in the three magnetic elements are definitely and simply related to one another. The electric current system responsible for these variations (cf. § 16) is illustrated in Fig. 7 of the paper cited; the

* *Proc. Roy. Soc.*, A, 95, p. 61, 1918.

† Viewed from the sun, over the right-hand hemisphere of the earth the local time is afternoon, or P.M., and this is therefore termed the P.M. hemisphere; the left-hand is similarly referred to as the A.M. hemisphere, while the sunlit and dark halves of the globe are called the day and night hemispheres.

*general* electric current system introduced by the storm, and responsible for the regular storm variations as a whole, is shown in Fig. 1. The currents are more intense and crowded together over the P.M. than over the A.M. hemisphere.

The irregular perturbations which are the most striking features of storms usually cease before the horizontal magnetic field has recovered its normal value, and before the disturbance

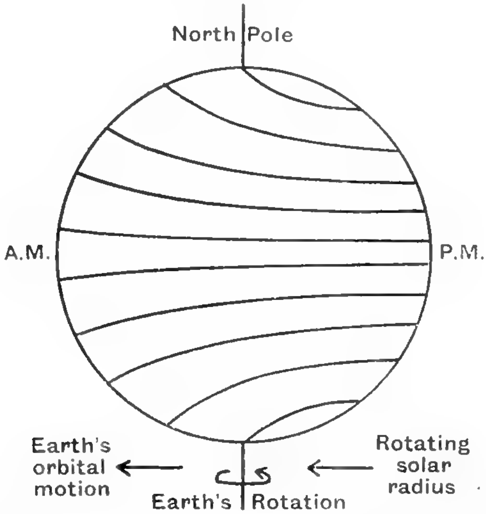



Fig. 1. The general distribution of additional electric currents in the earth's atmosphere during a magnetic storm, as viewed from the sun.

diurnal variations have quite died away. Hence some tendency to an increase in the horizontal magnetic force, and some small residual disturbance diurnal variation, will appear on quiet days shortly succeeding a storm, although no disturbing causes may be acting at the time.

§3. Great magnetic disturbance, such as was dealt with in my paper on magnetic storms, does not occur by any means so often as, on the average, once a month. But disturbance effects of the kind there described characterize also the less disturbed states common on ordinary days, and in the polar regions are practically always existent. This implies that the principal storm effects may be traced in the difference between magnetic phenomena on ordinary and on the international quiet days. The mean horizontal force on the former is slightly less than on the latter, and the differences between the diurnal magnetic variations on ordinary (or all) days and on quiet days are similar in type to the disturbance diurnal variations during storms, though, of course, much less in amplitude*.

It has been mentioned that irregular perturbations during storms are most intense in polar regions, and that disturbance is scarcely ever absent in these localities. The disturbance diurnal variations are also specially marked there, as is shown by Fig. 2. This is derived from Dr Chree's discussion of the Antarctic magnetic results obtained by the British Expedition of 1911–12†, and

* I hope to trace this similarity in some detail in a future paper, at the same time considering the seasonal changes in the disturbance diurnal variations. In latitudes 50° to 70° these seasonal changes are of considerable interest.

† C. Chree, "Seventh Kelvin Lecture," *Journ. Inst. Elec. Eng.*, 54, pp. 405–425, 1915.

illustrates the difference between the all-day and quiet-day diurnal variations at Kew and in the Antarctic. The vector diagrams represent the changes of force in the horizontal plane, corrected for the non-periodic change*. The annual mean curves for all and quiet days are given separately. The Kew quiet-day curve is derived from the five international quiet days per month. The larger

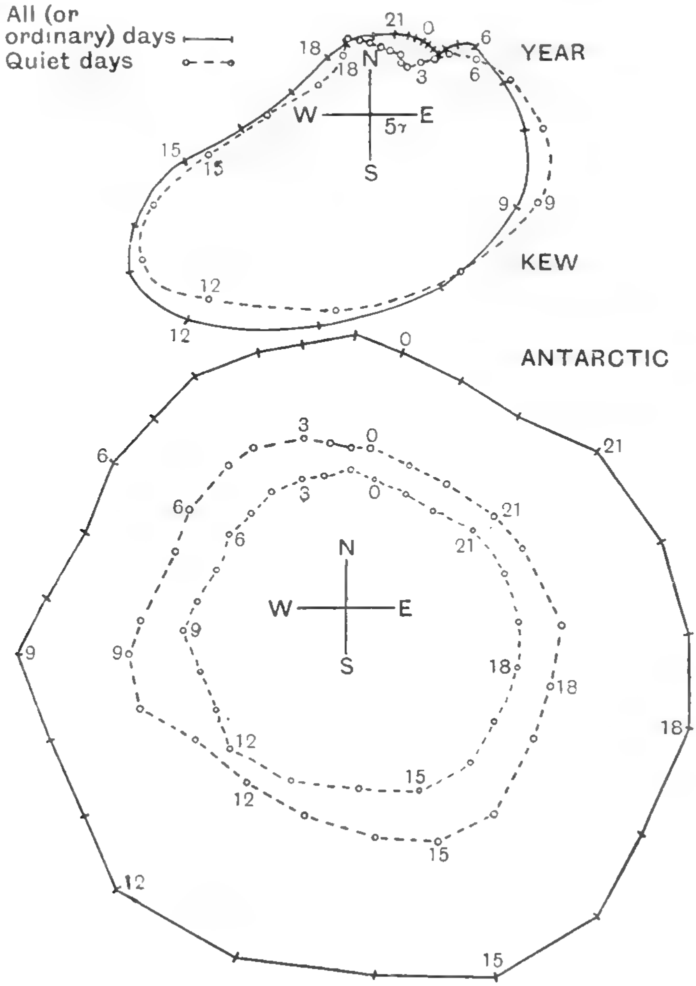



Fig. 2. Vector diagrams for the Kew and Antarctic annual mean diurnal variations of horizontal magnetic force for all (or ordinary) days and quiet days (at Kew 5, in the Antarctic 10, and 5 per month).

of the Antarctic quiet-day curves is obtained from ten days per month chosen by Dr Chree; later, for the winter season, he considered the five international days, and it appeared that the corresponding diurnal variation was practically a half-scale replica of the all-day variation†; I

* The horizontal magnetic force, being a vector, may be represented by a line $OP$ drawn with definite length and direction from a fixed point $O$. As the horizontal intensity and declination vary, the changes are indicated by the motion of $P$. When the changes which are not periodic in the course of a day are abstracted, the diurnal variation is represented by a closed curve, known as the vector diagram. This is described by $P$ at a varying rate, roughly shown by marking the points arrived at at different hours. The variations are small compared with the whole magnetic force, so that the origin $O$ cannot be shown on the diagrams.

† C. Chree, *l.c.*, Figs. 5, 6, 7.

have therefore drawn such a half-scale curve to represent the annual mean Antarctic curve which is to be compared with the quiet-day Kew curve.

The effect of disturbance on the diurnal variation is clearly far greater in the Antarctic than at Kew. Moreover, the similarity of type between the all and quiet-day curves in the former case indicates that what produces the difference between them, *i.e.* the disturbance diurnal variation, is probably the cause of the major part of the diurnal variation even on the five quietest days of a month. We may infer that the regular diurnal variation, free from the effect of disturbance, is very small in the Antarctic, and, in particular, a good deal less than that in the latitude of Kew or Greenwich.

§4. The chief general features of magnetic disturbance at any station depend upon the position of the station with respect to the earth's axes and to the sun. Both geographical and magnetic axes seem to have a share in determining the course of the phenomena, but in tropical and temperate latitudes the distinction between them is unimportant. Disturbance effects depend, both as regards incidence or frequency, and as regards intensity, on latitude, season, and local time. Widespread disturbance affects regions in high latitudes and over the P.M. hemisphere most strongly and with most irregular variability. As regards smaller and more local disturbance, similar preferences are found to exist, though varying somewhat with the special type of the disturbance. One type is the 'bay,' in which a simple perturbation takes place from and back to the normal intensity and direction of the field. These occur at all hours of the day and night, but more particularly between noon and midnight, *i.e.* over the P.M. hemisphere. Fig. 3 shows

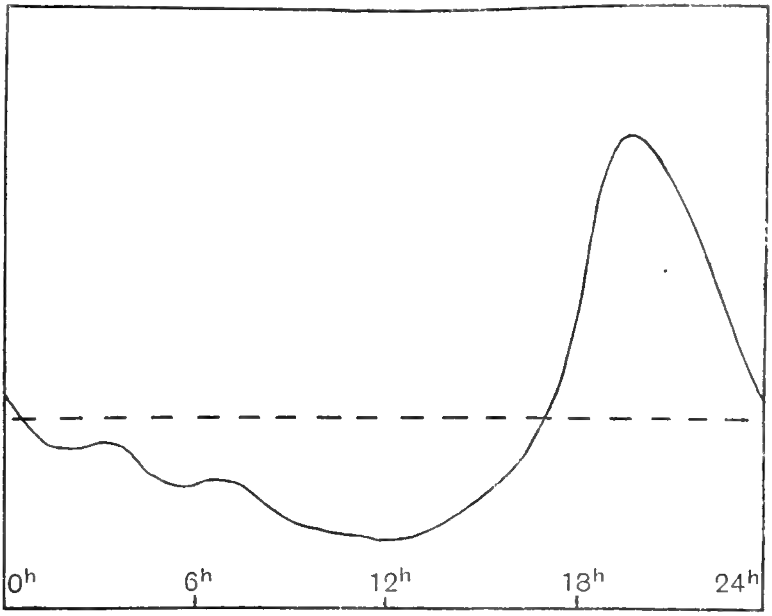



Fig. 3. Diurnal variation of frequency of occurrence of magnetic bays at Zikawei, relative to the mean number per hour (represented by the broken line).

the mean diurnal variation of frequency of such bays at Zikawei (1877–1908). Father de Moidrey has considered also other special types, such as pulsations and sudden twitches, from the same standpoint. These likewise occur more often after than before noon, but the epoch of maximum frequency occurs nearer midnight than at $18^h$, the night hemisphere being more favoured than either the day or the P.M. hemisphere*. Similar results are found at other stations in tropical and mean latitudes.

Disturbances of all kinds, both small and great, vary in frequency according to the time of year. W. Ellis found from the Greenwich records that they occur more often about the time of equinox than at the solstices, and the records of other observatories confirm this. Father Cortie† has suggested that this is due not to the position of the sun with respect to the earth's equator, but to the variation in the heliographic latitude of the earth; the earth is in the semi-equatorial plane at times near the equinoxes. This hypothesis seems not improbable in the light of the general conclusions of this discussion.

## THE REGULAR DAILY MAGNETIC VARIATIONS.

§ 5. The regular diurnal variations corresponding to ideally quiet magnetic conditions show a dependence on local time, latitude and season which is strikingly different from that displayed by magnetic disturbance. The characteristic relations with these three factors in the case of the intensity of the quiet-day changes may be summed up in one simple generalization: their intensity at any station at any time depends mainly on the zenith distance of the sun. It is greatest at stations then situated immediately 'beneath' the sun, it diminishes towards the twilight circle, and is small over the night hemisphere.

The *type* of the variations depends less simply on latitude and local time. There are, indeed, two sets of regular diurnal variations, which in many respects are very similar; one, the solar diurnal variation, is related to the sun's hour angle, and the other, the lunar diurnal variation, is related to the hour angle of the moon (in each case reckoned in 24 hours from one lower transit to the next). The two variations can easily be separated from one another. The mean (solar) hourly inequality is first derived, using all days in a lunation. Each lunar hour will have occurred with practically equal frequency at every solar hour, and the solar diurnal variation is thus free from lunar effects. When this is subtracted from the hourly values of the magnetic element, the residuals display the lunar influence by itself, or mixed up with disturbance variations having no connection with the moon. By rearranging the hourly values according to lunar time, the lunar diurnal magnetic variation is obtained.

§ 6. In the mean of a lunation the latter variation is found to be purely semidiurnal in each element, but this is by no means the case at any particular phase of the lunation, as is clearly shown by the corresponding vector diagram. The semidiurnal character of the *monthly mean* lunar diurnal variations causes the vector diagram to be an ellipse, described twice daily by the moving point $P$. But at each phase of the lunation the motion of $P$ is accelerated from about the time of sunrise: it remains greater than in the monthly mean curve, throughout the hours of sunlight, and diminishes at about sunset to less than the average motion, its course during the hours of darkness being quite short. The *lunar* hour at which sunrise occurs varies throughout the lunation, retrograding through 24 hours from $6^h$ at one new moon, through $0^h$ at first quarter, $18^h$ at full moon, and

* J. de Moidrey, *Terrestrial Magnetism*, 22, p. 39, 1917; also the volume of Zikawei magnetic observations for 1911.

† A. L. Cortie, *Monthly Notices R. A. S.*, 73, p. 58, 1912; 76, p. 13, 1916.

$12^h$ at third quarter, to $6^h$ at the succeeding new moon. Hence the magnification of the lunar changes at sunrise commences at all stages of the lunar day, and though the sun thus has an important influence on the intensity of the changes, their direction is chiefly governed by the

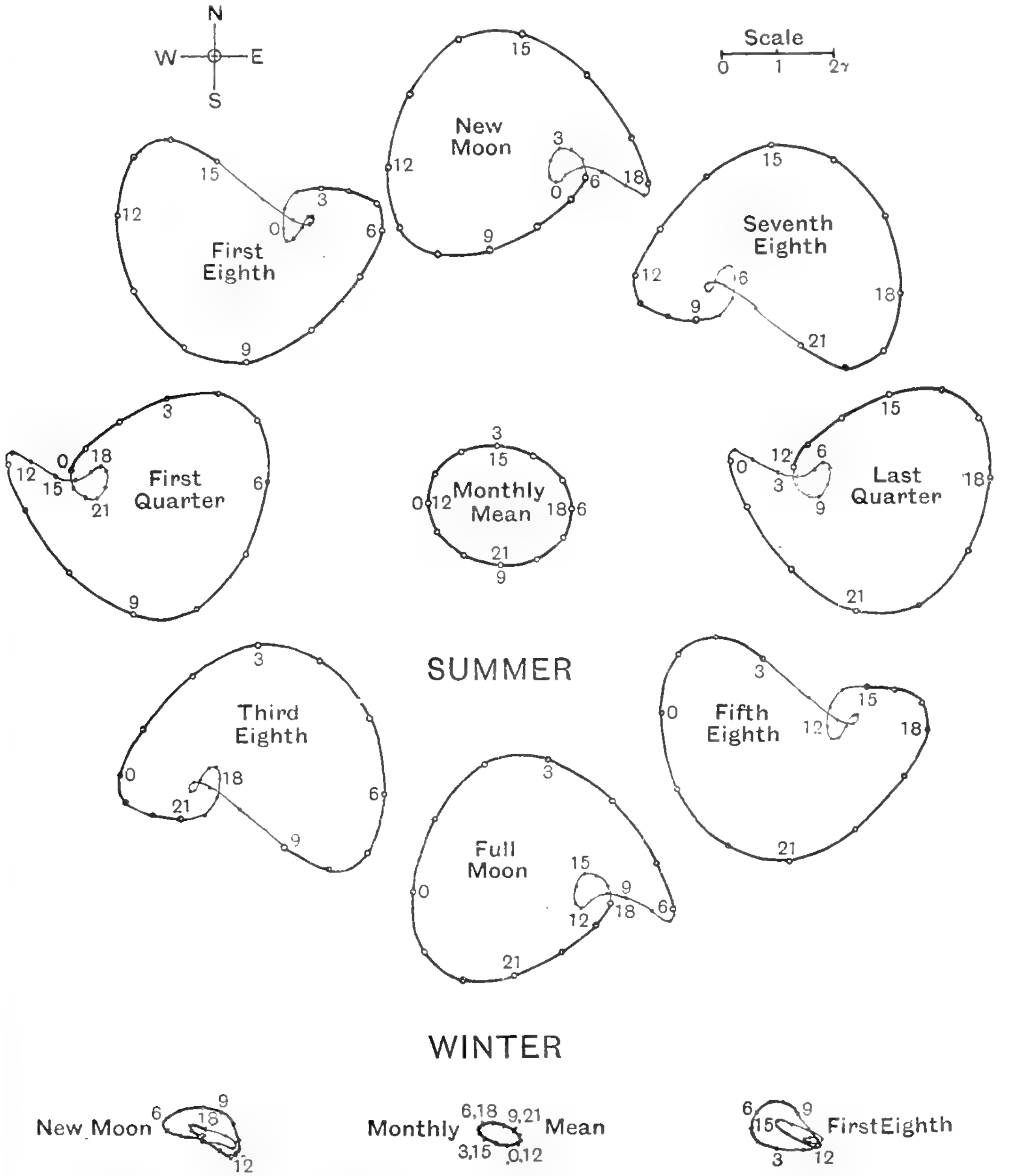



Fig. 4. Vector diagrams for the lunar diurnal variation of horizontal magnetic force at Pavlovsk, in summer (above) and winter (below): for the mean of a number of whole lunations (in the centre), and also at various particular epochs in the lunation.

moon. This is evident on comparing the direction of description of the vector diagrams in Fig. 4 at sunrise (say) at any lunar phase with the direction in the monthly mean diagram at the same lunar hour. Fig. 4 relates to the lunar diurnal variations of horizontal magnetic force at Pavlovsk (60° N.) for the two seasons of summer (May to August) and winter (November to February). The monthly mean diagram is in the centre, in each case; for summer eight diagrams are given for the separate phases of the moon*. These evidently comprise only two distinct curves, so that two only are drawn for the season of winter. The diagrams for the separate phases indicate, by the thickness of the lines, the parts corresponding to the two lunar half-days centred at solar midday and midnight. These roughly correspond with the periods of sunlight and darkness, which are, however, rather longer or shorter according to the season. At Pavlovsk the midday altitude of the sun is 56° at midsummer, and 7° in midwinter. There is but little difference between the 'day' and 'night' halves of the vector diagram at Pavlovsk in winter, either half being comparable with the night half in summer. The day portion in summer is, on the contrary, much enlarged. Fig. 4 consequently illustrates the influence of the sun's zenith distance as depending both on local (solar) time and on season.

§7. In the case of the solar diurnal variations the sun's zenith distance displays a similar relation to the intensity of the variations, though the demonstration is slightly complicated by the fact that here the sun also plays the part, taken in the former instance by the moon, of governing the type or direction of variation. In Fig. 5 are shown the vector diagrams for the quiet-day

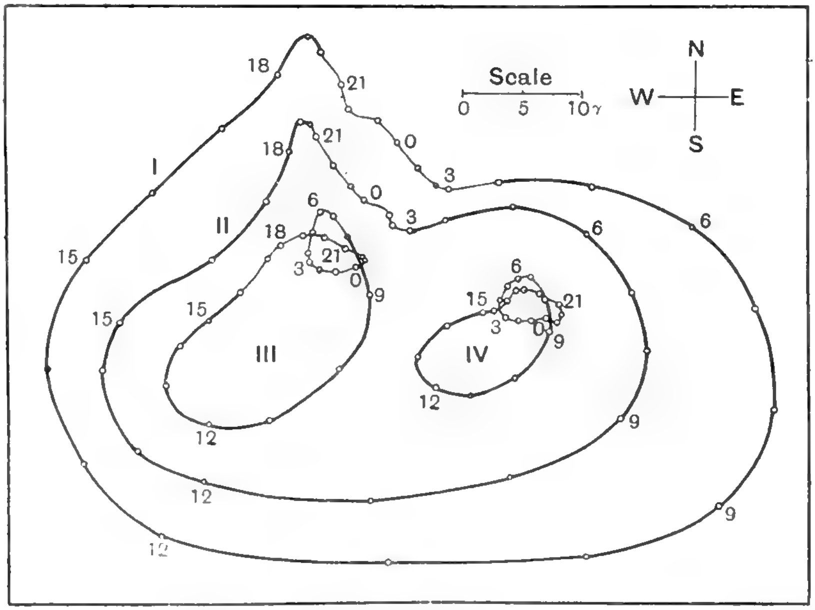



Fig. 5. Quiet-day vector diagrams of the daily variation of magnetic force in the horizontal plane at Greenwich, 1889–1914.

I. June, sunspot maximum years.
II. June, sunspot minimum years.
III. December, sunspot maximum years.
IV. December, sunspot minimum years.

* They are calculated from the Fourier coefficients given in Table VI A of my paper in the *Philosophical Transactions*, A, 214, p. 316, 1914; the phase angles are varied with the lunar phase as described on p. 301 of that paper. The use of the Fourier coefficients in this way is equivalent to smoothing the curves actually computed from the observations. The lunar variation is so minute that this is desirable for illustrative purposes.

The above variations relate to 'quiet' solar years, 1897–1903, during which the horizontal force and declination at Pavlovsk were approximately

$16550\gamma$ ($1\gamma = 10^{-5}$ c.g.s.) and 0°·6 E.

solar diurnal variations at Greenwich (51° N.) during the months of June (I, II) and December (III, IV), the two sets of curves, I, III and II, IV, being derived from different groups of years. As before, the day and night portions are drawn with different thicknesses, and the influence of the sun's zenith distance, as affected by local time and season, is seen to be similar to that indicated by Fig. 4. The seasons at Greenwich are, of course, rather less extreme than those at Pavlovsk.

§8. The sun's zenith distance also depends on the latitude of a station, and curves might be given to illustrate the dependence of the regular diurnal variations on this factor. This is well shown by the horizontal intensity variations, but the variations of declination vanish and change sign at the equator, so that the vector diagrams do not show an *areal* enlargement in tropical latitudes. But away from the equator, where this vanishing of the declination variation ceases to influence the curves, the vector diagrams would serve to illustrate the point. It will suffice, however, to refer to the remarks on Fig. 2 at the end of §3, to render it clear that the solar diurnal variations diminish greatly from temperate to high latitudes; the same is doubtless true also of the lunar variations, though there are no available polar data in their case.

## SOLAR FREQUENCIES AND TERRESTRIAL MAGNETIC CHANGES.

§9. So far the discussion has dealt with the manner in which magnetic disturbance phenomena or quiet-day variations are distributed *over the earth*, in relation to the earth's axes and the sun's geocentric coordinates. Another kind of dependence demands consideration, however, in which attention is directed to the earth's magnetic condition as a whole in connection with the succession *in time* of certain variable characteristics of the sun. These characteristics relate to the physical condition of the sun, which changes intrinsically, and also in its presentation to the earth, on account of the solar rotation. The existence of the solar influences already described would of itself suggest that any periodicities observable in the sun might appear also in the phenomena of the earth's magnetism, as indeed is found to be the case.

The two principal periods connected with the sun are that of the great solar cycle of activity indicated by sunspots, prominences, faculae, and so on, and the period of solar rotation. The former, which is an intrinsic period, is of somewhat variable duration, the average length being about 11 years. The latter period, relative to the earth, is approximately 27·3 days; the true rotation period is 25·2 days, for the sunspot zones, the difference of 2·1 days being due to the earth's orbital motion, which is in the same direction as that of *radii vectores* from the sun.

The solar cycle of activity affects the sun's surface as a whole, spots, faculae and prominences being indications of local disturbances which are symptomatic of the general condition of the sun. These local phenomena appear sporadically and irregularly, and last for a limited period of variable duration. But their average frequency and distribution in latitude vary in unison throughout the solar cycle. While this indicates that the visible agitation on the sun's disc is a consequence of changes of the surface in general, it is true that particular regions often remain abnormally subject to local outbreaks, not necessarily continuous, throughout several rotation periods*.

The rotation period is of importance to the earth's magnetic condition only in so far as there

* E. W. Maunder, *Monthly Notices R. A. S.*, 65, p. 555, 1905.

are inequalities in the physical state of the solar surface. Magnetic phenomena which depend on the sun's visible surface as a whole should show no relation to this period. Again, magnetic variations which are irregular in occurrence and intensity would naturally be associated with similar solar characteristics, so far as the former are found to be influenced by the sun.

The general march of the sun's activity as a whole throughout the solar cycle is well represented by the annual mean values of the spotted area ($A$) of the sun's visible hemisphere, as determined at Greenwich, or by Wolf's annual sunspot numbers ($S$) which measure the annual mean frequency of sunspots. The areas or numbers for individual days or months show considerable variation about the mean values $A$ or $S$.

§ 10. The relation between the general solar activity and the regular diurnal magnetic variations (§ 1) may be examined by comparing the range $R$ of the annual mean inequality for any magnetic element and station with the sunspot number $S$. The correspondence between them is remarkably close, and can be represented with considerable accuracy by a linear relation $R = a + bS$ (Wolf's formula), where $a$ and $b$ are constants for a particular station and element. The variation in range is unaccompanied by any marked change of type, as is illustrated by Fig. 5; the curves I, III relate to ten years of considerable, and II, IV to ten years of small, solar activity*. The two June and the two December curves are mutually similar, though their sizes are considerably different. These curves, being derived from quiet days (§ 7), represent very approximately the regular diurnal variations. Except in polar regions, moreover, the difference in range between the quiet and all day diurnal magnetic variations is not large (Fig. 2), so that the latter also agree with Wolf's formula, at tropical and temperate stations. The existing data for the *lunar* diurnal variations do not suffice to show whether these can likewise be represented by the formula $R = a + bS$.

§ 11. As regards the irregular magnetic changes, *i.e.* disturbances, Sabine found that the average frequency of magnetic *storms* shows a marked correspondence with Wolf's sunspot numbers $S$. But the relationship is less exact than in the case of Wolf's formula for the regular diurnal magnetic variations (§ 10). Years of few sunspots are (on the whole) conspicuously quiet magnetically, but years of like sunspot development have shown notably different degrees of magnetic disturbance. Dr Chree has instanced 1893 as a year which signally failed to show an amount of disturbance corresponding to the spottedness of the sun at the time†.

§ 12. The question now arises, How far do the above sunspot relationships indicate a connection with the general, and how far with the local, conditions of the sun's surface? The answer must depend on the extent to which the local sporadic solar disturbances appear to affect the various terrestrial magnetic phenomena.

The foregoing review suggests that magnetically 'quiet' or 'undisturbed' conditions correspond to the absence, or smallness, of some positive factor associated with 'disturbance.' At times of magnetic disturbance there are additions to the regular diurnal variations, these additions being of simple definite type and of amplitude depending on the degree of disturbance or—since they

* The ten year groups were 1891–5, 1905–9 and 1889, 1890, 1899–1902, 1911–14, the corresponding mean Greenwich values of spotted area (corrected for foreshortening) being 994 and 71, expressed in millionths of the area of the sun's visible hemisphere. The curves in Fig. 5 are derived from the quiet-day inequalities published in the Greenwich volumes, allowance being made for the non-periodic change.

† C. Chree, "Kelvin Lecture," p. 415.

die away only gradually after disturbance has ceased*—on the time elapsed since this cessation. The regular diurnal variations, which are of quite another type in relation to local time and terrestrial distribution, persist throughout with scarcely any change. In tropical and temperate latitudes the additional disturbance diurnal variations are small in comparison with the regular variations, except during magnetic storms. They are much less, indeed, in those latitudes, than the changes in the regular variations themselves from maximum to minimum epochs of solar activity (cf. the curves for Kew and Greenwich in Figs. 2, 5). It would seem that the regular diurnal variations vary gradually throughout the solar cycle, in unison with the *general* activity of the sun. The quiet-day variations at sunspot maximum or minimum retain amplitudes appropriate to these epochs, independently of the contemporary presence or absence of visible solar disturbance.

The secular march of magnetic disturbance shows no such regular pace. Magnetic storms—the most violent form of disturbance—sometimes break out at the very 'trough' of the solar cycle, though at such times, so far as I am aware, only when the sun also shows some sign of special local activity. This corresponds with the irregular occurrence of solar outbreaks, which sometimes appear for a brief period with local intensity at minimum epoch. The properties of extreme variability and intermittency are common to solar and to terrestrial magnetic 'disturbances,' and this correspondence, together with the 27·3-day periodicity described in § 13, suggests that magnetic disturbance is connected with local rather than with the general surface activity on the sun.

Not every local solar disturbance, however, finds its counterpart in the magnetic variations on the earth, nor is it yet possible to refer given magnetic disturbances to particular solar outbreaks shown as spots. The features of solar activity which are associated with spots or other visible phenomena may not be those which are most closely connected with the production of magnetic disturbance.

§ 13. There are few facts of greater significance, with respect to the relation between magnetic changes and the sun, than the tendency shown by the earth's magnetic activity to return to its condition at any particular time, after the lapse of one or more periods of synodic rotation of the sun, *i.e.* after intervals which are integral multiples of about 27·3 days. Great magnetic storms illustrate this tendency very clearly, for the number of pairs or series among them, in which the members are separated by such time intervals, is out of all proportion to that which would be expected on account merely of chance†. The time intervening between two storms can often not be determined to within a few hours, but the observed intervals often conform closely to the sun's synodic rotation period. This period varies for different belts of latitude on the sun; the rotation period of 27·3 days, which roughly agrees with the recurrence period in magnetic disturbance, is that of the belt in which sunspots are of most frequent occurrence. Since sunspots have proper motions of their own upon the sun, the rotation period has to be determined as an average from many spots; the range of variation shown by values from individual spots is of the same order as the range in the intervals between magnetic storms which are separated approximately by 27 days, or a multiple of it. The latter qualification is necessary because, just as there may be more than one disturbed region on the sun throughout a given period,

* Like the subsiding 'swell' after a storm on the ocean.

† E. W. Maunder, *Monthly Notices R. A. S.*, 65, pp. 2, 538, 666, 1904; 76, p. 63; F. W. Dyson, *Observatory*, 28, p. 176.

so there may be more than one series of storm recurrences simultaneously proceeding upon the earth. Also, just as intermittent outbreaks may occur in a particular solar area through several rotation periods, so magnetic storms sometimes show recurrences throughout a similar period, with gaps at one or more intervening 'recurrence' epochs. The recurrence phenomenon is correctly described as a *tendency*, affording some expectation of a recurrence of a storm or lesser disturbance, about 27·3 days after a given disturbance; a storm at one epoch will, however, no more be *inevitably* followed by another after this interval than will a sunspot be inevitably observable on a second presentation of its region to the earth after a rotation period.

The recurrence tendency is shown not only by great disturbances; Dr Chree, using daily character figures such as were mentioned in § 1, has demonstrated very clearly that disturbed and quiet days in general (with character figures 2 and 0) manifest the same tendency. The mean character figures for days round about the 27th, 54th, and 82nd day after a day of character (disturbed or quiet) much diverging from the average show a marked but diminishing tendency towards a repetition of this character*.

§ 14. The most convincing interpretation of the recurrence tendency shown by magnetic storms was given by Mr Maunder (*l.c.*), who independently discovered the phenomenon, which had been previously noted by Broun and others. He concluded that magnetic storms must be consequences of the presence in the earth's neighbourhood of some agency arising from a restricted area of the sun's surface, and travelling outwards in a limited stream in some particular direction. Such streams as are suitably directed will traverse the space in the earth's neighbourhood, overtaking the earth in its orbit on the P.M. side, as indicated in Fig. 1. Should the emitting area remain active over a sufficient period, projecting the stream nearly in the same direction throughout (relative to the solar surface), it may again traverse the space round the earth, after an interval of one or more rotations. In this way there may arise a recurrence tendency with the observed period.

Dr Chree's results may be explained along similar lines, if magnetic disturbance in general is referred to the agency of more or less well-defined streams emitted from particular disturbed localities on the solar surface. The manifestation of the recurrence tendency by quiet days is not to be regarded as due to the calming influence of limited solar areas, but simply to the absence of disturbing causes. When, as the sun rotates, the streams which it emits are projected so as to impinge upon and traverse the earth, magnetic disturbance of greater or less intensity results: when such streams happen to be absent from the space round the earth, magnetically quiet conditions prevail.

The recurrence tendency indicates that the solar regions which emit the streams often remain active and approximately stationary on the sun for one or more rotation periods. The constituents of the streams are projected with a speed sufficiently great to prevent any serious dissipation or sideways diffusion within a distance equal to the radius of the earth's orbit. Owing to the continual renewal of the streams from the emitting areas, the whole set of streams (if there are several existent at the same time) will appear to rotate with the sun, like the curved spokes of a wheel. There will be a certain lag of the streams, the curvature at any distance depending on the longitudinal and transverse speeds of the constituents; this angular lag, being probably

* C. Chree, *Phil. Trans.*, A, 212, p. 75, 1912; 213, p. 245, 1913; also the "Kelvin Lecture," *l.c.*

nearly the same for all streams in the plane of the ecliptic, will not affect the recurrence tendency. The sun's angular velocity is such that the transverse velocity of a stream, relative to the earth, is approximately $4 . 10^7$ cm. per sec., or about one-thousandth of the velocity of light. If the constituents of the stream take 24 hours to travel from sun to earth, the mean longitudinal velocity would be about four times that transverse velocity; if only one hour, about 100 times

The earth's angular diameter as viewed from the sun is very small (17"·6), so that any given stream-line would cross its 'solid' diameter in 35 seconds; it is therefore not difficult to understand why suddenly-commencing magnetic storms seem to start almost simultaneously over the whole earth. Some idea of the breadth of the intense streams concerned may be gained from the duration of storms. A duration of one day (which is not uncommon) would correspond to a breadth, in the ecliptic plane, of about 35 million kilometres, or an angular breadth, viewed from the sun, of about 13°. The sudden commencement which characterizes all very intense storms suggests that intense solar streams are somewhat sharply defined, at least on the forward side.

The general distribution of streams in the space round the sun is radial, though with a slight lag or curvature; the differences of intensity of the streams in different directions may be considerable. The radial distribution does not depend upon the streams being projected normally to the sun's surface: at several diameters' distance, the parallax due to oblique projection will be small. Owing to the varying situation of disturbed regions on the sun's surface, and to the varying directions of projection, many streams must miss the earth and so fail to produce any changes in the earth's magnetic field. The non-recurrence of particular stream-transits and storms may be due either to change in the direction of projection, to intermission of activity at the source, or perhaps to changes in the earth's heliographic latitude. Ignorance of the direction of projection from particular solar areas at any time precludes the possibility, at present, of identifying the precise solar region which is the source of given magnetic disturbances. It seems not unlikely, however, that on the whole sunspots or disturbed areas would be most effective when they are situated in or near the heliographic latitude of the earth. Father Cortie has urged that to this cause is due the observed seasonal inequality of disturbance frequency (§ 4).

§ 15. The main points which have so far emerged from this review of magnetic phenomena may be recapitulated as follows:

(*a*) Quiet magnetic conditions correspond to the absence from the earth's neighbourhood of certain solar emissions which, arising from locally disturbed solar areas, and being projected into space along confined streams, produce magnetic disturbance when they come into proximity with the earth. Disturbance phenomena are due to additional magnetic fields intermittently superposed upon the earth's permanent field and upon the field responsible for the regular diurnal variations.

(*b*) Disturbance effects are experienced with special intensity or frequency over particular regions of the earth, relative to the earth's axes and the geocentric coordinates of the sun. Polar regions, and the P.M. hemisphere, are those most affected.

(*c*) The regular diurnal magnetic variations are of definite type, quite different from that of the disturbance diurnal variations; they depend as to type on locality and season. The intensity of the changes proceeding at any station at any hour, so far as concerns these regular diurnal variations, is controlled by a solar agency other than that referred to in (*a*). This agency varies gradually in intensity throughout the solar cycle, and is connected with the general state of the sun rather than with localized outbreaks of activity. Its terrestrial magnetic effect is almost wholly confined to the day hemisphere of the earth, being greatest at any time over the regions immediately 'beneath' the sun. Its influence diminishes rapidly with increasing zenith distance, towards the twilight circle.

§ 16. Using the method of harmonic analysis by which Gauss proved that the source of the earth's main magnetic field is internal, Schuster has demonstrated the *external* origin of the regular solar diurnal magnetic variations*. The same conclusion has been shown to apply to the lunar diurnal variations†. It is true also of magnetic disturbance phenomena: the regular storm effects are proved to be of external origin by the sign of the vertical force variations, which would be reversed if the contrary hypothesis were true.

Attempts have been made in the past to explain the daily magnetic variations by changes in the magnetic permeability of the atmosphere, but without success. Electromagnetic action, arising from external electric currents, seems the only alternative. These currents might be supposed to flow either in the atmosphere or in the 'empty' space beyond. In the latter case the currents would presumably be supposed to consist of streams of electric particles from the sun, influencing the earth's magnetism by their own magnetic field. The *lunar* diurnal magnetic variations could hardly be accounted for in this way, and for this and other reasons the regular (solar and lunar) diurnal changes may best be attributed to currents flowing *within* the atmosphere. The same holds good for the magnetic disturbance currents, as is suggested by their close connection with aurorae, which are undeniably atmospheric electrical phenomena. The earth's aerial envelope is so relatively shallow that the electrical currents concerned in the production of the magnetic variations may be regarded as forming approximately spherical current sheets: that is, they flow in nearly horizontal atmospheric layers.

The current density over these sheets can be calculated, with little uncertainty and without any extraneous hypothesis, from the observed magnetic changes. But the thickness of the layer, and its situation, are not known, so that the current density per unit *area* of cross-section cannot be determined.

The currents flow in layers of definite electrical conductivity under the impulsion of certain electromotive forces. If the latter can be independently determined, with the aid of the known current density it becomes possible to calculate the electrical conductivity of the layer as a whole; that of unit thickness of the layer can be deduced only when the total depth has been ascertained.

§ 17. The electromotive forces responsible for the regular diurnal magnetic variations can be independently calculated by the use of a hypothesis which seems firmly grounded. The solar and lunar diurnal variations of the barometer indicate the existence of two world-wide atmospheric circulations with periods of a solar and a lunar day. The motion is almost entirely horizontal, and nearly symmetrical with respect to the equator. Such a horizontal movement of

* A. Schuster, *Phil. Trans.*, A, 180, p. 467, 1889, and 208, p. 163, 1907. The primary external field induces also a secondary internal field of varying magnetic force.

† van Bemmelen, *Met. Zeitschrift*, p. 218, 1912; p. 589, 1913; also S. Chapman, *Phil. Trans.*, A, 218, p. 1, 1919.

air will induce *horizontal* electromotive forces (which alone will be of importance for the flow of the electric currents described) in conjunction with the vertical component of the earth's main magnetic field. As I have remarked elsewhere*, the magnetic variations indicate that the actual electric current systems have precisely those properties of symmetry or reversed symmetry, with respect to the equator, which would be possessed by the systems of electromotive forces produced in the manner supposed. This confirmation of the hypothesis is strengthened by the agreement in period between the lunar diurnal atmospheric circulation and the monthly mean lunar diurnal variations, both of which are purely semidiurnal. The solar diurnal atmospheric circulation seems likewise to be mainly semidiurnal; the corresponding magnetic variations also have important components of other periods, but this is readily to be accounted for by the independent solar agency, which controls the intensity of the magnetic variations (§ 15 (*c*)). The same feature is shown by the lunar magnetic changes at any particular phase of the lunation (§ 6, Fig. 4), although there the lunar directing influence is clearly semidiurnal. By mathematical analysis it is possible to infer the functional dependence of this local control factor upon the sun's zenith distance, using the lunar diurnal magnetic variations only; this relation being known, it is found that the observed type of solar diurnal magnetic variations can be explained, in a general way, as a consequence of the atmospheric circulation indicated by the barometric variations.

The solar and lunar diurnal barometric variations depend mainly on the movements of the lower atmospheric strata, since these contain most of the total mass of air. But since the lunar diurnal circulation can hardly be other than the effect of tidal forces, which everywhere act proportionally to the mass attracted, the tidal motion may be expected to extend throughout the whole atmosphere, or, at any rate, to a great height; ultimately the increasing kinematic viscosity of the air will reduce its amplitude. The solar semidiurnal atmospheric circulation likewise appears to be a very fundamental oscillation probably affecting all strata up to a high level, though perhaps with modifications of phase. If, therefore, the horizontal motion is nearly the same at all levels, the electromotive forces induced will be likewise similar, and are calculable from the known magnitudes of the barometric variations and the earth's vertical magnetic force. By identifying these electromotive forces with those responsible for the electric currents deduced from the diurnal magnetic variations (the identification being justified on the ground of similarity of type) it is possible to determine the electrical conductivity of the current sheet (§ 16).

§ 18. The magnitude of the conductivity, as first deduced in this manner by Schuster†, was remarkably high, and later estimates have increased it rather than otherwise‡. By assuming an outside limit of 300 kilometres for the thickness of the layer, Schuster calculated a lower limit for the mean *specific* conductivity. This was much greater than the specific conductivity in the lower atmosphere; the electromotive forces induced here cannot contribute appreciably to the diurnal magnetic variations. The current sheet must consequently be situated higher up, probably at a high level in the stratosphere, where the air is greatly rarefied.

The distribution of conductivity being found clearly dependent on the sun's zenith distance, the conclusion at once suggests itself that the sun exercises its control over the intensity of

* *Observatory*, p. 52, Jan., 1918.

† A. Schuster, *Phil. Trans.*, A, 208, p. 181, § 13, 1907.

‡ S. Chapman, *Phil. Trans.*, A, 218, p. 1, 1919.

the regular diurnal magnetic variations by determining the conductivity of the air in the higher layers. In other words, the 'general' controlling agency propagated from the sun, which affects mainly the day hemisphere of the earth (§ 15 (*c*)), acts by ionizing the air, probably in some fairly definite layer.

§ 19. The nature of the solar ionizing agent was discussed by Schuster in the paper previously referred to. He concluded that the only possible alternatives were ultra-violet light, and ions projected from the sun with sufficient speed to generate new ions in the atmosphere by impact. At the time there was no clear evidence as to whether ultra-violet waves were capable of ionizing dust-free air; it was remarked also that "it seems difficult to believe that, even if emitted by the hottest portions of the sun's envelope, they are not absorbed again by the surrounding cooler layers." Nevertheless the view was retained, as a possibility, "that the powerful ionization of the air, which we must consider to be an established fact, is a direct effect of solar radiation."

More recently Swann* has discussed the possibility of accounting for an atmospheric conductivity corresponding to Schuster's estimate, by ultra-violet radiation of the amount which would fall upon the earth's atmosphere if the sun radiates approximately like a 'black body.' He concluded that the amount was quite inadequate for the purpose.

§ 20. The review of the two types of magnetic variations considered in §§ 1–15 led to the conclusion that the solar agent which is concerned in the production of magnetic disturbance is distinct from that which controls the intensity of the regular diurnal variations. The latter acts by ionizing the atmosphere: the question arises, What is the nature and action of the former solar agent?

As regards its nature, it is immediately possible to assert, negatively, that it cannot be ultra-violet light, since this would affect the day hemisphere almost exclusively, whereas the disturbance agent favours the P.M. or night hemisphere. Neither can it consist of merely material particles without electric charge, for there is no reason why these should crowd towards the polar regions; in other ways, also, these are incompatible with, and unable to account for, several features of magnetic disturbance. Electric particles of some kind offer the only alternative, and such may well be supposed to issue from the locally disturbed regions of the sun with which the disturbance agent has been associated in the course of the previous discussion. Strong electric fields, which might supply the necessary energy of projection, would appear to exist on the sun's surface, for though no Stark effect has been observed, Prof. Strutt has shown that solar prominence movements can hardly be accounted for except by strong electrostatic forces†. Moreover, electric corpuscles projected into the neighbourhood of the earth will suffer deflection in the earth's magnetic field; a right and left-hand asymmetry of distribution is by no means unlikely, and the experiments of Birkeland and the calculations of Störmer upon the paths of corpuscles projected towards a uniformly magnetized sphere show that the particles would tend to be deflected towards the magnetic poles‡. The investigation of the paths is a

* W. F. G. Swann, *Terrestrial Magnetism*, 21, p. 1, 1916.

† Cf. Deslandres, *Comptes Rendus*, 155, p. 1579, 1912; also Strutt, *Monthly Notices R. A. S.*, 77, p. 65, 1916.

‡ An excellent bibliography of the researches of Birkeland and Störmer is given by Vegard in "Nordlichtuntersuchungen," Kristiania, *Vid. Sk.* I, *Mat. Natur. Kl.*, 1916.

matter of great analytical difficulty, but there seems to be good prospect of accounting for the favoured regions of magnetic disturbance along these lines. It certainly seems to be demonstrable that the particles might be so deflected as to bend round the earth and fall upon the hemisphere which is invisible from the sun. The undoubted connection between aurora and magnetic disturbance also affords strong support to the view that magnetic disturbance is originated by electric corpuscles from the sun.

If the particles are projected by strong electrostatic fields in the emitting areas, the members of any one stream will be of like charge, as regards sign. Possibly the particles of different streams may be of opposite kinds, but at present I know of no magnetic evidence which suggests this. In the case of any one stream, however, the injection of numbers of like charges into the atmosphere may be expected to have two effects. One effect would be an increased ionization and electrical conductivity of the air, perhaps with considerable local inequalities. These inequalities would persist for some time, but the injected charge would tend to distribute itself uniformly over the atmosphere, under the influence of the electric forces arising from the inequalities of surface density of charge. This tendency would rapidly take effect, owing to the conductivity produced in the layer. At the same time the electrified air would tend to expand, owing to the mutual repulsion of the imprisoned charge; escape being possible only by upward motion, the layer would expand upwards, until the charge was gradually dissipated—part of the air being carried away with it. This vertical motion would induce horizontal electromotive forces in the charged, ionized layer, in conjunction with the horizontal component of the earth's magnetic field. In a recent paper on magnetic storms I have shown that these electrical forces are such as would account for the magnetic variations associated with storms. The case in which the injection of corpuscles is local, rather than world-wide, has not been considered in detail as yet, but probably 'bays' (§ 4) are produced in this way. It seems likely that in all types of magnetic disturbance the solar electric corpuscles are the cause both of the electromotive forces and the ionization involved in the production and maintenance of the atmospheric electric currents.

**§ 21.** These conclusions regarding the 'disturbance' solar agent have a direct bearing on the 'general' solar agent which affects the regular diurnal magnetic variations over the sunlit hemisphere. If the former consists of electrical corpuscles, the latter cannot do so—no mere difference of mass or sign of charge would account for the complete difference of distribution of the two agents on reaching the earth. On the other hand, the apparently sole alternative among possible ionizing agents, viz. ultra-violet light, seems to accord with all the properties which the 'general' solar agent has been shown to possess: for the latter affects the sunlit hemisphere almost exclusively, it arises from the sun's surface as a whole, and its intensity varies only gradually, from time to time, in correspondence with the general activity of the sun. The identification seems incontrovertible, despite the difficulty of understanding how such radiation escapes from the solar atmosphere; this difficulty, and the apparently great intensity which it must have to account for the high degree of atmospheric ionization produced by it, only add to the interest of the conclusion, for solar physics. The regular and large variation of intensity throughout the solar cycle is also of much interest. The range of variation in the conductivity of the current-sheet, deduced from the magnetic variations, is from 30 to 50 per cent., according to the particular solar cycle. According to Swann (*l.c.*), the ionizing radiation should vary in intensity as the square of the conductivity, so that its range must be of the order of 100 per cent. This short wave radiation wholly fails to penetrate to the earth's surface; the ordinary solar spectrum stops short at about λ2900, and beyond this wave-length the total intensity of radiation undergoes no variation comparable with 100 per cent. The observations of Abbot, Fowle, and Aldrich* suggest that the 'solar constant' varies from time to time by small amounts, rarely rising to 10 per cent., but these variations occur irregularly, and so far as I am aware there appears to be no definite evidence for a systematic change following the course of the solar cycle.

It is difficult to believe that if the ultra-violet part of the sun's spectrum is merely radiation of incandescence, it is so unconnected with the remaining portion as to suffer large changes without some parallel variation in the longer-wave spectrum. There is little likelihood that the spectrum of the solar radiation before entry into the earth's atmosphere is of black-body type, but it may be of interest to remark on the relative changes of intensity in the different parts of the spectrum on the supposition that these correspond merely to changes of a black-body spectrum of varying temperature. The bases of the calculation will be the same as those chosen by Swann in his discussion of atmospheric ionization (§ 19), viz., a mean solar temperature of 6000° absolute, and the hypothesis that only the radiation of wave-length less than λ1350 is concerned in ionization. A variation of 100 per cent. in this extreme section of the spectrum would correspond to a variation of 230° in the temperature. This would involve a shifting of the wave-length of maximum intensity from about λ4700 to λ4520, and a change in the total intensity of radiation for the whole spectrum amounting to 13 per cent. The change in the visible spectrum would be rather smaller than this.

§ 22. The facts hitherto reviewed may next be considered in their bearing upon atmospheric questions. One such question is, Are the layers affected by the two kinds of solar emissions the same or different, and, if different, what is their relative situation?

Even *a priori* it would be expected that two such different emissions as corpuscles and aether-waves will have different powers of penetration into the atmosphere, though it would not be possible, on such grounds alone, to decide whether the 'absorbing' layers were wholly distinct or not. The magnetic phenomena, however, give a fairly clear indication that they are practically distinct without overlapping. If both ionizing influences were experienced in the same layer, the conductivity of the latter should be much increased at times of great disturbance; the result should resemble that produced by the augmented ultra-violet radiation at sunspot maximum, *i.e.* there should be a general magnification of the regular diurnal magnetic variations. This should apply to all the harmonic components, though not necessarily equally, since the disturbance ionization is not distributed in the same way as the regular ionization. The observed results do not agree with this; some changes occur in the first harmonic component, and the production of these has been independently accounted for†; but the other components remain the same or, possibly, are slightly diminished.

On the alternative hypothesis that there are two distinct layers of ionization, some increase in the ordinary diurnal magnetic variations at times of disturbance would still be expected, unless it is supposed that the atmospheric circulation (§ 17) present in the greater part of the atmosphere, including the layer ionized by ultra-violet light, does not exist in the other ionized layer. From this conclusion, which seems unavoidable, follows the further inference that the

* C. G. Abbot, *Proc. Nat. Acad. Sci.*, 1, p. 331, 1915.

† *Proc. Roy. Soc.*, A, 98, p. 61, 1918.

magnetic-disturbance layer is situated at a higher level than the diurnal-variations layer. It seems justifiable to identify the former with the stratum in which aurorae are observed (though the disturbance currents are not to be supposed confined to the latitudes in which these luminous phenomena are visible). The lower limit of the auroral layer is fairly definite, being at a height of about 90 kilometres*. Aurorae sometimes extend to a height of 300 kilometres, but the great majority seem to occur between 90 and 120 kilometres.

The diurnal-variations layer is presumably situated somewhere between the top of the troposphere (10 km.) and the base of the auroral layer (90 km.). Its extreme thickness can therefore not exceed 80 kilometres, and on this account the lower limit of the specific electrical conductivity of this layer, deduced by Schuster on the basis of a limit of thickness of 300 kilometres, must be increased fourfold. On this and other grounds described elsewhere† the original estimate of $10^{-13}$ c.g.s. should be increased to about $3 \cdot 10^{-12}$ for points immediately beneath the sun, at sunspot maximum. This, moreover, is still only a lower limiting value.

§ 23. Profs. Fowler and Strutt‡ have recently confirmed the truth of a suggestion made many years ago by Hartley, to the effect that the limitation of the solar spectrum at the violet end is due to absorption by ozone in the earth's atmosphere. Prof. Strutt§ has also demonstrated that this ozone does not exist in sufficient amount in the lower atmosphere; it must be in the stratosphere, where alone permanent differences of composition in different layers are possible. It may be assumed that the ozone is produced in the layer which is ionized by ultra-violet light, and possibly the presence of ozone in that layer may afford escape from the difficulties which appeared in Swann's calculations.

It is a matter for observational enquiry to determine whether ozone is produced also in the auroral layer, ionized by electric corpuscles. If so, at times of magnetic disturbance such ozone would probably intercept part of the ultra-violet radiation before it reached the layers where, by its own ionizing action, and by the presence of atmospheric diurnal circulations, the diurnal magnetic variations are produced. Perhaps to this cause may be ascribed the slight diminution which seems to occur in the regular magnetic variations at such times.

NOTE (added July, 1919). In a paper read (on May 22, 1919) before the Institution of Electrical Engineers, and shortly to be published, I have suggested that the ultra-violet radiation discussion in § 21 may be some type of $\gamma$-radiation, and that the corpuscles are (as Vegard has urged) $\alpha$-particles. If both these originate from radio-active processes on the sun, the $\gamma$-rays would be expected to penetrate more deeply into our atmosphere than the $\alpha$-particles—which agrees with the relative situation of the two absorbing layers, according to the inference drawn in § 20.

Since this paper was written, I have discovered that the lunar-diurnal magnetic variation, unlike the corresponding solar-diurnal changes, varies considerably in amplitude according to the degree of magnetic disturbance existing at the time. This suggests that the lunar-diurnal component of the atmospheric circulation extends upwards beyond the range of the solar-diurnal component into the auroral layer. These new data will be described and discussed in a later paper.

* The auroral researches of Birkeland, Störmer, Vegard, Krogness and others are summarized and the references catalogued by Vegard in the memoir referred to on p. 356, footnote.

† Cf. the third footnote on p. 355.

‡ A. Fowler and R. J. Strutt, *Proc. Roy. Soc.*, A, 93, p. 577, 1917.

§ R. J. Strutt, *ibid.*, A, 94, p. 260, 1917.

# XVII. *On the Representations of a Number as a Sum of an Odd Number of Squares.*

By L. J. MORDELL, B.A. (Cantab.), Birkbeck College, London.

[*Communicated by* Mr G. H. Hardy. *Read* 3 February 1919.]

The first results concerning the number of representations of a given number as a sum of an odd number of squares were given by Gauss* for the case of three squares, and were found from the arithmetical theory of the ternary quadratic. Eisenstein† followed with some results in the cases of five and seven squares, publishing them without proof. H. J. S. Smith‡ completed these results and published them also without proof. About ten years after the publication of Smith's paper, the proof of Eisenstein's results was set as a prize problem by the French Academy of Sciences. Papers were submitted by Smith§ and Minkowski‖, who were awarded equal prizes. Their proofs¶ depended upon the arithmetical theory of the general quadratic form, and are examples of some of the most delicate and intricate demonstrations to be found in the whole range of mathematical analysis.

The results for the representations of a number as a sum of an even number of squares may also be proved by means of non-arithmetical principles depending on the expansions of Elliptic functions. The results for three squares had also been proved by means of class relation formulae, a method** involving the expansion of products and quotients of theta functions, and also some delicate points concerning the expression of the non-equivalent binary quadratics of a given determinant. In the case of the other odd numbers of squares, the problem, until just recently, proved intractable to analytic methods.

A general method, however, for dealing with such questions was recently developed by the author††. The whole question of finding the number of representations of any number $n$ as a sum of $r$ squares is equivalent to finding the expansion of $\theta_{00}^r$, where as usual, if $q = e^{\pi i \omega}$,

$$\theta_{00} = 1 + 2q + 2q^4 + 2q^9 + \dots.$$

From the theory of the modular functions, it follows that $\theta_{00}^r$ can be expressed as the sum of a finite number of functions called modular invariants. These can be expressed in a variety of ways, *e.g.* as products of various theta functions, or as power series in $q$ wherein the coefficients depend on the representations of $n$ as a number of squares less than $r$. When $r$ is even, there is

* *Disquisitiones Arithmeticae*, article 291.

† "Note sur la représentation d'un nombre par la somme de cinq carrés," *Crelle's Journal*, vol. XXXV. p. 368.

‡ "On the orders and genera of quadratic forms containing more than three indeterminates," *Collected Math. Papers*, vol. I. p. 521.

§ "Mémoire sur la représentation des nombres par des sommes de cinq carrés," *Collected Math. Papers*, vol. II. p. 623.

‖ "Mémoire sur la théorie des formes quadratiques à coefficients entières," *Gesammelte Abhandlungen*, vol. I. p. 3.

¶ See also Bachmann, *Zahlentheorie*, vol. IV.

** This method was discovered independently by Kronecker and Hermite. See also my papers "On class relation formulae," *Messenger of Mathematics*, vol. XLVI. p. 113; "Note on class relation formulae," *Messenger of Mathematics*, vol. XLV. p. 76.

†† "On the representations of numbers as a sum of $2r$ squares," *Quarterly Journal of Pure and Applied Mathematics*, vol. XLVIII. p. 93.

no difficulty in finding for one of the invariants a function which possesses a simple expansion in powers of $q$, and which gives the well-known results when $r=2$, 4, 6 or 8, as only one invariant is required in these cases. For this invariant, say $\chi$, a doubly infinite series of the type $\sum_a \sum_b (a+b\omega)^{-s}$ at once suggests itself, and from its form is seen to be an invariant of the type required for the expansion of $\theta_{00}^{2s}$. The series is then converted into a singly infinite series in powers of $q=e^{\pi i\omega}$.

While the same method applies whether $r$ is odd or even, in the first case I was unable to find a doubly infinite series for $\chi$, and so could not complete the solution. Mr Hardy, however, in the course of his investigations on asymptotic formulae for the number of representations, was led to a series which, by means of the principles given in my paper, he showed was the required invariant $\chi$. As soon as I saw Mr Hardy's work* for the case of five squares, which he communicated to me in various letters, the true origin of the expression for the $\chi$ invariant occurred to me. From the work which follows it will be seen that we can now solve the question of finding the number of representations of a number by any definite quadratic form.

A few preliminary considerations may make this paper more intelligible to the reader. In addition, the general method given in my previous paper will be presented in a more simple form.

It is well known that, putting

$$\omega' = \frac{c+d\omega}{a+b\omega},$$

where $a$, $b$, $c$ and $d$ are integers satisfying the equation

$$ad-bc=1,$$

and also the congruence

$$\begin{vmatrix} a, & b \\ c, & d \end{vmatrix} \equiv \begin{vmatrix} 1, & 0 \\ 0, & 1 \end{vmatrix} \text{ or } \begin{vmatrix} 0, & -1 \\ 1, & 0 \end{vmatrix} \pmod 2,$$

then

$$\theta_{00}(0, \omega') = \xi\sqrt{(a+b\omega)}\,\theta_{00}(0, \omega),$$

where $\xi$ depends on $a, b, c, d$ only and not on $\omega$, and a convention is made whereby a definite one of the values of the radical is taken. The substitutions by which $\omega$ is changed to $\omega'$ form a group denoted by $\Gamma_3$, which is generated by the substitutions $\omega'=\omega+2$ (denoted by $S^2$) and $\omega'=-1/\omega$ (denoted by $T$). Writing $\omega=x+iy$, the fundamental polygon of this group is the part of the upper $\omega$ plane bounded by the lines $x=\pm 1$ and the circle $x^2+y^2=1$, but only the right-hand half of the boundary is reckoned as belonging to the fundamental polygon.

Suppose now that we have linearly independent functions of $\omega$, say $\psi_1, \psi_2, \psi_3, \ldots$, which are such that, first, $\psi_\lambda/\theta_{00}^r$ is an automorphic function for the substitutions of the group $\Gamma_3$; this will be the case if it is unchanged by the substitutions $S^2$ and $T$. Secondly, suppose that $\psi_\lambda/\theta_{00}^r$ has no singularities within the fundamental polygon except when the denominator is zero†. This occurs at $\omega=1$, and putting $\omega=1-1/\Omega$, $Q=e^{\pi i\Omega}$, so that $\omega=1$ corresponds to

* See also his paper "On the representation of a number as the sum of any number of squares, and in particular of five or seven," *Proceedings of the National Academy of Sciences* (Washington, U.S.A.), vol. IV. p. 189.

† To study the behaviour of $\theta_{00}(\omega)$ at the rational point $\omega=c/a$, write

$$\omega=\frac{c\Omega+d}{a\Omega+b},$$

$\Omega = i\infty$, the expansion of $\theta_{00}{}^r$ starts with $Q^{r/4}$, if we neglect unimportant factors. Finally, suppose that no additional singularity at $\omega = 1$ arises from the numerator $\psi_\lambda$, which we assume can be expressed as an ordinary power series in $q$, say

$$\psi_\lambda = \lambda_0 + \lambda_1 q + \lambda_2 q^2 + \dots.$$

We now introduce the modular function $\Delta(\omega)$ defined by

$$\Delta(\omega) = q^2 \prod_1^\infty (1 - q^{2n})^{24},$$

and consider a function $f(\omega)$ defined by

$$\Delta^r(\omega) f(\omega) = (\theta_{00}{}^r - A_1\psi_1 - A_2\psi_2 - \dots - A_\lambda\psi_\lambda)^{24},$$

wherein $A_1, A_2, \dots A_\lambda$ are constants. Since

$$\theta_{00}(\omega+2) = \theta_{00}(\omega), \quad \theta_{00}(-1/\omega) = \sqrt{(-i\omega)}\,\theta_{00},$$
$$\Delta(\omega+2) = \Delta(\omega), \quad \Delta(-1/\omega) = \omega^{12}\Delta(\omega),$$

$f(\omega)$ is an automorphic function for the group $\Gamma_3$. Now $\Delta(\omega)$ has three zeros within the fundamental polygon, namely two at infinity (*i.e.* at $\omega = i\infty$) and one at $\omega = 1$, so that the order of the poles of $f(\omega)$ is $3r$. The constants $A_1, A_2, \dots A_\lambda$ can be taken so that the expansion of the numerator of $f(\omega)$ starts with $q^{24\lambda}$, that is, the order of the zeros of $f(\omega)$ is at least $24\lambda$. Hence if $24\lambda > 3r$, or say

$$\lambda = 1 + I(\tfrac{1}{8}r),$$

where $I(\frac{1}{8}r)$ denotes the integral part of $\frac{1}{8}r$, $f(\omega)$ has more zeros than poles inside the fundamental polygon and hence vanishes identically. Therefore

$$\theta_{00}{}^r = A_1\psi_1 + A_2\psi_2 + \dots + A_\lambda\psi_\lambda.$$

But $\lambda - 1$ of these invariants $\psi$ are given by the functions

$$\chi_t = \theta_{00}{}^r \left(\frac{\theta_{01}{}^2\theta_{10}{}^2}{\theta_{00}{}^4}\right)^{2t},$$

where $$t = 1, 2, \dots I(\tfrac{1}{8}r).$$

For it is clear that, first, $\chi_t/\theta_{00}{}^r$ is unaltered by the substitutions $S^2$ and $T$; secondly, the functions $\chi_t$ are linearly independent; and finally, no extra singularities are introduced in the function $\chi_t/\theta_{00}{}^r$, since $r - 8t \geqslant 0$. Hence we require only one additional invariant, say $\chi$, to complete the solution.

Mr Hardy's work suggested to me the following form* for $\chi$,

$$\chi = \Sigma\left[\frac{\theta_{00}(\omega)}{\theta_{00}\left(\dfrac{c+d\omega}{a+b\omega}\right)}\right]^r,$$

so that $\Omega = i\infty$ corresponds to $\omega = c/a$. We then have

$$\theta_{00}(\omega) = \xi\sqrt{(a\Omega + b)}\,\theta_{g,h}(\Omega),$$

where $\xi$ is a constant and $g$, $h$ depend on $a$, $b$, $c$ and $d$. The expansion of $\theta_{g,h}(\Omega)$ in powers of $Q = e^{\pi i\Omega}$ then gives the required result. The factor $\sqrt{(a\Omega+b)}$ has no influence on the result, as it will be cancelled by a corresponding factor in the numerator.

* The corresponding invariant for the representations of a number by any definite quadratic form, say $f(x, y, z, \dots)$ with $r$ variables, would be $\phi(\omega)/\phi(\omega')$, where

$$\phi(\omega) = \Sigma\Sigma\Sigma e^{\pi i\omega f(x, y, z, \dots)}.$$

The substitutions

$$\omega' = \frac{c+d\omega}{a+b\omega}$$

are such that

$$\phi(\omega)/\phi(\omega') = \xi(a+b\omega)^{\frac{1}{2}r}$$

with $\xi$ independent of $\omega$, and are well known.

where the summation refers to all the substitutions of the group $\Gamma_3$, except that, for a given pair of values for $a$ and $b$, we take only one set of values of $c$ and $d$, as the general term of the series is independent of the values of $c$ and $d$. Further, we need not consider negative values of $b$; and if $b=0$, we take only $a=d=1$, so that each term of the series occurs once only. From the congruence satisfied by $a$, $b$, $c$ and $d$, it is clear that the function $\chi/\theta_{00}{}^r$ is unaltered by the substitutions $S^2$ and $T$, which merely alter the order of the terms in $\chi$, an absolutely convergent series when $r>4$, as is evident from the next few lines.

From the well-known formula for the linear transformation of the theta functions, it follows* that

$$\chi = 1 + \sum_{a,b}\left[\frac{e^{-\pi i a b^2 c/4} H_{a,b}}{\sqrt{[-ib(a+b\omega)]}}\right]^r,$$

where the square root is taken with a positive real part. The summation refers to all coprime values of $a$ and $b$ of opposite parity, except that $b$ is always positive and $b=0$ is excluded from he summation.

Also
$$H_{a,b} = \sum_{s=0}^{b-1} e^{-\pi i a (s-\frac{1}{2}b)^2/b}$$
$$= i^{-\frac{1}{2}a}\sqrt{b}\left(\frac{b}{a}\right) \quad \text{(if } a \text{ is odd)}$$
$$= i^{-\frac{1}{2}(a-1)(b-1)-\frac{1}{2}a}\sqrt{b}\left(\frac{a}{b}\right) \quad \text{(if } b \text{ is odd)},$$

where $\left(\frac{b}{a}\right)$ and $\left(\frac{a}{b}\right)$ are the Legendre-Jacobi symbols of quadratic reciprocity.

By considering separately the cases when $a$ is odd or even, we find that

$$\chi = 1 + \sum_{a,b}\left[\frac{H}{\sqrt{[-ib(a+b\bar{\omega})]}}\right]^r,$$

where
$$H = \sum_{s=0}^{b-1} e^{-\pi i a s^2/b}$$
$$= i^{-\frac{1}{2}a}\sqrt{b}\left(\frac{b}{a}\right) \quad \text{(if } a \text{ is odd)}$$
$$= i^{\frac{1}{2}(b-1)}\sqrt{b}\left(\frac{a}{b}\right) \quad \text{(if } b \text{ is odd)}.$$

We must now convert $\chi$ into a singly infinite series. If $r$ is even, $\chi$ reduces to the doubly infinite series used in my previous paper†. Suppose‡ then that $r$ is odd and that

$$r \equiv \kappa \pmod 8.$$

Then, since $H/\sqrt{b}$ is an eighth root of unity, we have when $a$ is odd

$$\left(\frac{H}{\sqrt{b}}\right)^r = \frac{i^{-\frac{1}{2}a(\kappa-1)}H}{\sqrt{b}} = \frac{i^{-\frac{1}{2}a(\kappa-1)}}{\sqrt{b}}\sum_{s=0}^{b-1} e^{-\pi i a s^2/b},$$

* See for example H. J. S. Smith, "Memoir on the theta and omega functions," *Collected Math. Papers*, vol. II. p. 474. The result is of course due to Hermite.

† *Quarterly Journal*, vol. XLVIII. p. 93.

‡ This is based on Hardy's work for the case $r=5$, communicated to me by letter.

but when $a$ is even

$$\left(\frac{H}{\sqrt{b}}\right)^r = \frac{i^{\frac{1}{2}(b-1)(\kappa-1)}H}{\sqrt{b}} = \frac{i^{\frac{1}{2}(b-1)(\kappa-1)}}{\sqrt{b}}\sum_{s=0}^{b-1} e^{-\pi i a s^2/b}.$$

Hence*

$$\chi = 1 + \sum_{a,b}\left[\frac{\tau \sum_{s=0}^{b-1} e^{-\pi i a s^2/b}}{\sqrt{b}\,[-i(a+b\omega)]^{\frac{1}{2}r}}\right].$$

where

$$\tau = i^{-\frac{1}{2}a(\kappa-1)} \text{ or } i^{\frac{1}{2}(b-1)(\kappa-1)},$$

according as $a$ is odd or even. In this expression for $\chi$, $a$ and $b$ are prime to each other. We can remove this restriction as follows. Writing

$$S_{\frac{1}{2}(r-1)} = \frac{1}{1^{\frac{1}{2}(r-1)}} + \frac{1}{3^{\frac{1}{2}(r-1)}} + \frac{1}{5^{\frac{1}{2}(r-1)}} + \dots \quad (\text{if } r \equiv 1 \pmod 4)$$

$$= \frac{1}{1^{\frac{1}{2}(r-1)}} - \frac{1}{3^{\frac{1}{2}(r-1)}} + \frac{1}{5^{\frac{1}{2}(r-1)}} - \dots \quad (\text{if } r \equiv 3 \pmod 4)),$$

and multiplying throughout by the series for $\chi$, we find

$$S_{\frac{1}{2}(r-1)}\chi = S_{\frac{1}{2}(r-1)} + \sum_{a,b}\left[\frac{\tau \sum_{s=0}^{b-1} e^{-\pi i a s^2/b}}{\sqrt{b}\,[-i(a+b\omega)]^{\frac{1}{2}r}}\right].$$

where now $a$ and $b$ need no longer to be prime to each other, but are still of opposite parity.

This is easily proved, for take $r \equiv 3 \pmod 4$ say. If $a$ and $b$ have now a common factor $k$ (odd of course), put

$$a = kA, \quad b = kB,$$

then

$$\tau \sum_{s=0}^{b-1} e^{-\pi i a s^2/b} = i^{-\frac{1}{2}kA(\kappa-1)} \sum_{s=0}^{kB-1} e^{-\pi i A s^2/B} = (-1)^{\frac{1}{2}(k-1)} k\, i^{-\frac{1}{2}A(\kappa-1)} \sum_{s=0}^{B-1} e^{-\pi i A s^2/B}.$$

Also

$$\sqrt{b}\,[-i(a+b\omega)]^{\frac{1}{2}r} = k^{\frac{1}{2}(r+1)}\sqrt{B}\,[-i(A+B\omega)]^{\frac{1}{2}r},$$

so that the term in $\Sigma_{a,b}$ becomes the product of $(-1)^{\frac{1}{2}(k-1)}k^{-\frac{1}{2}(r-1)}$ by the corresponding term in $\Sigma_{A,B}$. A similar argument holds when $a$ is even because

$$i^{\frac{1}{2}(kB-1)(\kappa-1)} = (-1)^{\frac{1}{2}(k-1)} i^{\frac{1}{2}(B-1)(\kappa-1)},$$

where $\kappa \equiv 3 \pmod 4$.

Writing now $-a$ for $a$ in the last formula for $\chi$, we have

$$S_{\frac{1}{2}(r-1)}\chi = S_{\frac{1}{2}(r-1)} + \sum_{a,b}\left[\frac{\tau \sum_{s=0}^{b-1} e^{\pi i a s^2/b}}{\sqrt{b}\,[i(a-b\omega)]^{\frac{1}{2}r}}\right],$$

where

$$\tau = i^{\frac{1}{2}a(\kappa-1)} \text{ or } i^{\frac{1}{2}(b-1)(\kappa-1)},$$

according as $a$ is odd or even, and the summation refers to all integer values of opposite parity for $a$ and $b$, except that $b > 0$.

* It should be noted that $[-i(a+b\omega)]^{\frac{1}{2}r}$ stands for $\{\sqrt{[-i(a+b\omega)]}\}^r$, where the radical as usual is taken with a positive real part.

The summation with respect to $a$ can be carried out as follows. Take first the part of the series arising from odd values of $a$. Consider the contour integral

$$\frac{1}{2}\int \frac{e^{\pi i(\theta/b+(\kappa-1)/4)\xi}}{\sqrt{b}\,[i(\xi-b\omega)]^{\frac{1}{2}r}(e^{\pi i\xi}+1)}\,d\xi=\frac{1}{2}\int F d\xi,$$

say, where the contour consists of an infinite circle, centre at the origin, with a cross cut from the point $\infty\omega$ of the circle encircling the point $\xi=b\omega$, in a clockwise direction and so that all the points $\xi=\pm 1, \pm 3, \dots$ are within the contour. The radical is defined by

$$-\tfrac{1}{2}\pi<\arg\,[i(t-b\omega)]^{\frac{1}{2}}<\tfrac{1}{2}\pi.$$

The integral taken around the infinite circle is zero if

$$0\leqslant\theta/b+(\kappa-1)/4<1,$$

or say

$$-\tfrac{1}{4}b(\kappa-1)\leqslant\theta<-\tfrac{1}{4}b(\kappa-5).$$

Hence the integral taken around the cross cut is equal to $2\pi i$ times the sum of the residues of the integrand at the points $\xi=a$. This gives

$$\frac{1}{2}\int_{\infty\omega}^{(b\omega-)} F d\xi=-\sum_{a}\frac{e^{\pi i\theta a/b+\pi i(\kappa-1)a/4}}{\sqrt{b}\,[i(a-b\omega)]^{\frac{1}{2}r}},$$

the notation for the contour integral being that employed, for example, in Whittaker and Watson's *Modern Analysis*. If now in the last series for $\chi$ we put $s^2=\theta+bm$, where $\theta$ satisfies the above inequality (and we can always find an integer $m$ so that this is the case for a given value of $s$), then

$$e^{\pi i\theta a/b}=e^{\pi i s^2 a/b-\pi i m a}=(-1)^m e^{\pi i s^2 a/b}.$$

Hence when we sum for odd values of $a$, the general term of the series for $S_{\frac{1}{2}(r-1)}\chi$ can be replaced by the integral

$$-\tfrac{1}{2}(-1)^m\int_{\infty\omega}^{(b\omega-)}\frac{e^{\pi i(\theta/b+(\kappa-1)/4)\xi}\,d\xi}{\sqrt{b}[i(\xi-b\omega)]^{\frac{1}{2}r}(e^{\pi i\xi}+1)},$$

and as $a$ is odd, $b$ must be even.

Similarly by summing for even values of $a$, we obtain the integral

$$\tfrac{1}{2}\,i^{\frac{1}{2}(b-1)(\kappa-1)}\int_{\infty\omega}^{(b\omega-)}\frac{e^{\pi i\theta\xi/b}\,d\xi}{\sqrt{b}[i(\xi-b\omega)]^{\frac{1}{2}r}(e^{\pi i\xi}-1)}$$

since now

$$e^{\pi i\theta a/b}=e^{\pi i s^2 a/b-\pi i m a}=e^{\pi i s^2 a/b},$$

where* as before $s^2=\theta+bm$, but now $0\leqslant\theta<b$.

Hence we have

$$S_{\frac{1}{2}(r-1)}\chi-S_{\frac{1}{2}(r-1)}=\sum_{b\text{ odd}}\tfrac{1}{2}\,i^{\frac{1}{2}(b-1)(\kappa-1)}\int_{\infty\omega}^{(b\omega-)}\frac{e^{\pi i\theta\xi/b}\,d\xi}{\sqrt{b}[i(\xi-b\omega)]^{\frac{1}{2}r}(e^{\pi i\xi}-1)}$$

$$-\sum_{b\text{ even}}\tfrac{1}{2}(-1)^m\int_{\infty\omega}^{(b\omega-)}\frac{e^{\pi i(\theta/b+(\kappa-1)/4)\xi}\,d\xi}{\sqrt{b}[i(\xi-b\omega)]^{\frac{1}{2}r}(e^{\pi i\xi}+1)}.$$

* No confusion will arise from the fact that the $\theta$'s satisfy different inequalities.

Putting $\xi = b\omega + \zeta$ and $q = e^{\pi i \omega}$, this becomes

$$S_{\frac{1}{2}(r-1)}\chi - S_{\frac{1}{2}(r-1)} = \sum_{b\text{ odd}} \tfrac{1}{2} i^{\frac{1}{2}(b-1)(\kappa-1)} \int_{\infty\omega}^{(0-)} \frac{q^{\theta} e^{\pi i \theta\zeta/b}\, d\zeta}{\sqrt{b}\,(i\zeta)^{\frac{1}{2}r}(q^b e^{\pi i \zeta} - 1)}$$

$$- \sum_{b\text{ even}} \tfrac{1}{2}(-1)^m \int_{\infty\omega}^{(0-)} \frac{q^{\theta + b(\kappa-1)/4}\, e^{\pi i (\theta/b + (\kappa-1)/4)\zeta}\, d\zeta}{\sqrt{b}\,(i\zeta)^{\frac{1}{2}r}(q^b e^{\pi i\zeta} + 1)}.$$

Expanding the integrands as power series in $q$, the right-hand side of this equation becomes

$$-\tfrac{1}{2} \sum_{b\text{ odd}} \sum_{\sigma=0}^{\infty} \frac{i^{\frac{1}{2}(b-1)(\kappa-1)}\, q^{\theta+b\sigma}}{\sqrt{b}} \int_{\infty\omega}^{(0-)} \frac{e^{\pi i(\theta/b+\sigma)\zeta}}{(i\zeta)^{\frac{1}{2}r}}\, d\zeta$$

$$- \tfrac{1}{2} \sum_{b\text{ even}} \sum_{\sigma=0}^{\infty} \frac{(-1)^{m+\sigma}\, q^{\theta+b\sigma+b(\kappa-1)/4}}{\sqrt{b}} \int_{\infty\omega}^{(0-)} \frac{e^{\pi i(\theta/b+\sigma+(\kappa-1)/4)\zeta}}{(i\zeta)^{\frac{1}{2}r}}\, d\zeta.$$

These are Hankel's well-known integrals* which occur in the theory of the $\Gamma$ function, and so this expression becomes

$$\frac{\pi}{\Gamma(\frac{1}{2}r)} \sum_{b\text{ odd}} \sum_{\sigma=0}^{\infty} \frac{i^{\frac{1}{2}(b-1)(\kappa-1)}\, q^{\theta+b\sigma}}{\sqrt{b}} \left(\frac{\pi\theta}{b} + \pi\sigma\right)^{\frac{1}{2}r-1}$$

$$+ \frac{\pi}{\Gamma(\frac{1}{2}r)} \sum_{b\text{ even}} \sum_{\sigma=0}^{\infty} \frac{(-1)^{m+\sigma}\, q^{\theta+b\sigma+b(\kappa-1)/4}}{\sqrt{b}} \left(\frac{\pi\theta}{b} + \pi\sigma + \frac{\pi}{4}(\kappa-1)\right)^{\frac{1}{2}r-1},$$

which reduces to

$$\frac{\pi^{\frac{1}{2}r}}{\Gamma(\frac{1}{2}r)} \sum_{b\text{ odd}} \sum_{\sigma=0}^{\infty} \frac{i^{\frac{1}{2}(b-1)(\kappa-1)}\, q^{\theta+b\sigma} (\theta+b\sigma)^{\frac{1}{2}r-1}}{b^{\frac{1}{2}(r-1)}}$$

$$+ \frac{\pi^{\frac{1}{2}r}}{\Gamma(\frac{1}{2}r)} \sum_{b\text{ even}} \sum_{\sigma=0}^{\infty} \frac{(-1)^{m+\sigma} q^{\theta+b\sigma+b(\kappa-1)/4} (\theta+b\sigma+b(\kappa-1)/4)^{\frac{1}{2}r-1}}{b^{\frac{1}{2}(r-1)}}$$

These two series can be expressed in a more simple form as follows.

Putting in the first series $\theta + b\sigma = M$, so that

$$M = s^2 - bm + b\sigma \equiv s^2 \pmod{b},$$

it becomes

$$\sum_{b\text{ odd}} \sum_{\sigma=0}^{\infty} q^M M^{\frac{1}{2}r-1} i^{\frac{1}{2}(b-1)(\kappa-1)} / b^{\frac{1}{2}(r-1)} = \sum_{M=1}^{\infty} A_M M^{\frac{1}{2}r-1} q^M,$$

say, where

$$A_M = \sum_{b\text{ odd}} \frac{f(b)\, i^{\frac{1}{2}(b-1)(\kappa-1)}}{b^{\frac{1}{2}(r-1)}} = \sum_{b\text{ odd}} \frac{f(b)\, i^{\frac{1}{2}(b-1)(r-1)}}{b^{\frac{1}{2}(r-1)}}$$

and $f(b)$ denotes the number of solutions of the congruence

$$s^2 \equiv M \pmod{b}.$$

Similarly the second series, when we put $\theta + b\sigma + b(\kappa-1)/4 = M$, so that

$$M = s^2 - bm + b\sigma + b(\kappa-1)/4 \equiv s^2 + b(\kappa-1)/4 \pmod{b}$$

* See for example Whittaker and Watson, *Modern Analysis*, p. 239. Putting $\omega = in$ ($n$ real) for example, and $\xi = it$, the integrals reduce practically to the standard form

$$\frac{1}{\Gamma(z)} = \frac{i}{\pi} \int_{\infty}^{(0+)} (-t)^{-z} e^{-t}\, dt.$$

and also
$$(-1)^{m+\sigma} = (-1)^{(s^2 - M + b(\kappa-1)/4)/b},$$

becomes
$$\sum_{b \text{ even}} \sum_{\sigma=0}^{\infty} q^M M^{\frac{1}{2}r-1} (-1)^{m+\sigma} / b^{\frac{1}{2}(r-1)} = \sum_{M=1}^{\infty} B_M M^{\frac{1}{2}r-1} q^M,$$

where
$$B_M = \sum_{b \text{ even}} \frac{\phi(b)(-1)^{(s^2-M+b(\kappa-1)/4)/b}}{b^{\frac{1}{2}(r-1)}} = \sum_{b \text{ even}} \frac{\phi(b)(-1)^{(s^2-M+b(r-1)/4)/b}}{b^{\frac{1}{2}(r-1)}}$$

and $\phi(b)$ denotes the number of solutions of the congruence
$$s^2 \equiv M - b(r-1)/4 \pmod{b}.$$

We now have
$$S_{\frac{1}{2}(r-1)}\chi - S_{\frac{1}{2}(r-1)} = \frac{\pi^{\frac{1}{2}r}}{\Gamma(\frac{1}{2}r)} \left[ \sum_{M=1}^{\infty} (A_M + B_M) M^{\frac{1}{2}r-1} q^M \right],$$

and this gives the expansion of $\chi$ as a power series in $q$, and we note that the constant term is unity.

It only remains now to consider the effect of the function $\chi$ upon the singularities* of the function $\chi/\theta_{00}{}^r$, in the fundamental polygon. From the pseudo-automorphic character of the series for $\chi$, it is clear that singularities can be introduced only at the point $\omega = i\infty$, but the power series for $\chi$ shows that this is not the case.

The function $\chi$ is also linearly independent of the functions $\theta_{00}{}^{r-8t}\,\theta_{01}{}^{4t}\,\theta_{10}{}^{4t}$, whose expansion starts with $q^t$, and hence satisfies all the necessary conditions. Since the expansion of $\chi$ starts with $1 + aq + \ldots$, we have at once
$$\theta_{00}{}^r = \chi + \Sigma\, C_t \theta_{00}{}^{r-8t} \theta_{01}{}^{4t} \theta_{10}{}^{4t},$$

where $t = 1, 2, \ldots I(\frac{1}{8}r)$, and the constants $C_t$ can be found by equating the coefficients of powers of $q$ on both sides.

When $r = 5$ or $7$, this equation reduces to†
$$\theta_{00}{}^r = \chi.$$

A similar identity holds when $r = 3$, but the doubly infinite series for $\chi$ is then semiconvergent, and the various transformations require justification.

When‡ $r = 5$ the result can be written as
$$\theta_{00}{}^5 = 1 + \tfrac{32}{3} \sum_{M=1}^{\infty} (A_M + B_M) M^{\frac{3}{2}} q^M,$$

where
$$A_M = \sum_{b \text{ odd}} \frac{f(b)}{b^2}, \quad B_M = -\sum_{b \text{ even}} \frac{(-1)^{(s^2-M)/b} f(b)}{b^2},$$

and $f(b)$ is the number of solutions of the congruence
$$s^2 \equiv M \pmod{b}.$$

* There is no need to consider the point $\omega = 1$, for if we put $\omega = 1 - 1/\Omega$ then
$$\frac{\chi}{\theta_{00}{}^r} = \frac{F(\Omega)}{\theta_{10}{}^r(\Omega)},$$
say, where $F(\Omega)$ vanishes when $\Omega = i\infty$, and its expansion in powers of $Q = e^{\pi i \Omega}$ cannot contain any negative powers of $Q$.

† $\chi$ is of course different for different values of $r$, but no confusion can arise thereby.

‡ The solutions for five and seven squares are in the forms given by Mr Hardy.

When $r=7$, we have

$$\theta_{00}^7 = 1 + \tfrac{256}{15} \sum_{M=1}^{\infty} (A_M + B_M)\, M^{\frac{5}{2}} q^M,$$

where
$$A_M = \sum_{b\,\text{odd}} \frac{(-1)^{\frac{1}{2}(b-1)} f(b)}{b^3}, \quad B_M = \sum_{b\,\text{even}} \frac{(-1)^{(s^2-M-\frac{1}{2}b)/b}\, \phi(b)}{b^3},$$

and $\phi(b)$ is the number of solutions of the congruence

$$s^2 \equiv M + \tfrac{1}{2}b \pmod{b}.$$

The series of the types $A_M$ and $B_M$ can be reduced to the form $\sum_n \left(\frac{\pm M}{n}\right)\frac{1}{n^s}$, as given by Eisenstein in the case of five and seven squares. The work however is rather detailed* because of the number of cases that arise according to the form of $M$. By starting from the known results, we can also deduce the doubly infinite series for $\chi$. Similar methods may be applied to the series $\sum_{n=1}^{\infty} F(n)q^n$, where $F(n)$ is the number of uneven classes of binary quadratics of determinant $-n$, the classes derived from the form $(1, 0, 1)$ being reckoned as $\frac{1}{2}$. The starting-point of the investigation is the well-known formula for the number of properly primitive classes of determinant $-n$. This formula is one of the type by which we express the number of representations of a number as a sum of three squares. The generating function of the series $\Sigma F(n)q^n$ is an expression similar to that denoted by $\chi$, but I shall return to the subject in another paper.

When $r = 9$, 11, 13, or 15, we have an equation of the form

$$\theta_{00}^r = \chi + C_r \theta_{00}^{r-8} \theta_{01}^4 \theta_{10}^4.$$

Equating the coefficients of $q$ on both sides, we have

$$16\, C_r = 2r - \frac{\pi^{\frac{1}{2}r}(A_1 + B_1)}{\Gamma(\frac{1}{2}r)\, S_{\frac{1}{2}(r-1)}}.$$

Now
$$A_1 = \sum_{b\,\text{odd}} \frac{i^{\frac{1}{2}(b-1)(r-1)} f(b)}{b^{\frac{1}{2}(r-1)}},$$

where $f(b)$ is the number of solutions of the congruence

$$s^2 \equiv 1 \pmod{b}.$$

But if $p$ and $q$ are prime to each other

$$f(pq) = f(p) f(q),$$

and the same functional relation holds if $f(p)$ is replaced by $i^{\frac{1}{2}(b-1)(r-1)}/p^{\frac{1}{2}(r-1)}$. Hence $A_1$ can be expressed as the infinite product

$$\prod_p \left[1 + \sum_{n=1}^{\infty} \frac{i^{\frac{1}{2}(p^n-1)(r-1)} f(p^n)}{p^{\frac{1}{2}n(r-1)}}\right],$$

where $p$ refers to the odd primes 3, 5, 7, .... But $f(p^n) = 2$ and

$$\tfrac{1}{4}(p^n - 1)(r-1) \equiv \tfrac{1}{4}n(p-1)(r-1) \pmod{2},$$

so that the infinite product becomes

$$\prod_p \left[1 + \frac{2(-1)^{\frac{1}{4}(p-1)(r-1)} p^{-\frac{1}{2}(r-1)}}{1 - (-1)^{\frac{1}{4}(p-1)(r-1)} p^{-\frac{1}{2}(r-1)}}\right] = \prod_p \left[\frac{1 + (-1)^{\frac{1}{4}(p-1)(r-1)} p^{-\frac{1}{2}(r-1)}}{1 - (-1)^{\frac{1}{4}(p-1)(r-1)} p^{-\frac{1}{2}(r-1)}}\right].$$

* See for example De Seguier, *Formes quadratiques et multiplication complexe*, p. 60.

Multiplying numerator and denominator by the denominator, this product becomes

$$\left[\sum_{m \text{ odd}} (-1)^{\frac{1}{4}(m-1)(r-1)} m^{-\frac{1}{2}(r-1)}\right]^2 \Big/ \left[\sum_{m \text{ odd}} m^{-(r-1)}\right],$$

where now $m$ takes all odd positive integer values. Hence

$$\frac{\pi^{\frac{1}{2}r} A_1}{\Gamma(\frac{1}{2}r) S_{\frac{1}{2}(r-1)}} = \frac{\pi^{\frac{1}{2}r} \sum\limits_{m \text{ odd}} (-1)^{\frac{1}{4}(m-1)(r-1)} m^{-\frac{1}{2}(r-1)}}{\Gamma(\frac{1}{2}r) \sum\limits_{m \text{ odd}} m^{-(r-1)}},$$

and can be expressed in finite terms by Euler's or Bernoulli's numbers.

The working for $B_1$ is slightly different. For

$$B_1 = \sum_{b \text{ even}} \frac{(-1)^{(s^2-1+\frac{1}{4}b(r-1))/b} \phi(b)}{b^{\frac{1}{2}(r-1)}},$$

where $\phi(b)$ is the number of solutions of the congruence

$$s^2 \equiv 1 - \tfrac{1}{4}b(r-1) \pmod{b}.$$

We first discuss the case when $r \equiv 1 \pmod 4$, so that

$$B_1 = (-1)^{\frac{1}{4}(r-1)} \sum_{b \text{ even}} \frac{(-1)^{(s^2-1)/b} f(b)}{b^{\frac{1}{2}(r-1)}},$$

where $f(b)$ is, as before, the number of solutions of the congruence

$$s^2 \equiv 1 \pmod{b}.$$

In the series for $B_1$, we can ignore values of $b$ divisible by 8. For, if $b = 8\beta$, the solutions of

$$s^2 \equiv 1 \pmod{8\beta}$$

can be grouped in pairs such as $s$ and $4\beta - s$, for which

$$\frac{s^2-1}{8\beta} \equiv \frac{(s-4\beta)^2-1}{8\beta} + 1 \pmod 2.$$

If $b$ is twice or four times an odd number, we note that $s$ is odd, so that

$$s^2 \equiv 1 \pmod 8$$

and

$$(-1)^{(s^2-1)/b} = 1.$$

Hence we have

$$B_1 = (-1)^{\frac{1}{4}(r-1)} \left(\frac{1}{2^{\frac{1}{2}(r-1)}} + \frac{2}{4^{\frac{1}{2}(r-1)}}\right) \sum_{b \text{ odd}} \frac{f(b)}{b^{\frac{1}{2}(r-1)}}.$$

From the formulae for $A_1$, it easily follows that the part of the coefficient of $q$ in $\chi$ arising from $B_1$ can be written as

$$\frac{\pi^{\frac{1}{2}r} B_1}{\Gamma(\frac{1}{2}r) S_{\frac{1}{2}(r-1)}} = \frac{(-1)^{\frac{1}{4}(r-1)} \pi^{\frac{1}{2}r}}{\Gamma(\frac{1}{2}r)} \left(\frac{1}{2^{\frac{1}{2}(r-1)}} + \frac{2}{4^{\frac{1}{2}(r-1)}}\right) \frac{\sum\limits_{m \text{ odd}} m^{-\frac{1}{2}(r-1)}}{\sum\limits_{m \text{ odd}} m^{-(r-1)}}.$$

Discussing now the case when $r \equiv 3 \pmod 4$, we have

$$B_1 = \sum_{b \text{ even}} \frac{(-1)^{(s^2-1+\frac{1}{2}b)/b} \phi(b)}{b^{\frac{1}{2}(r-1)}}.$$

where $\phi(b)$ is the number of solutions of the congruence

$$s^2 \equiv 1 - \tfrac{1}{2}b \pmod{b}.$$

In the summation we ignore values of $b$ divisible by 16, for if $b = 16\beta$ the solutions of the congruence

$$s^2 \equiv 1 - 8\beta \pmod{16}$$

can be arranged in pairs such as $s$ and $8\beta - s$, for which

$$\frac{s^2 - 1 + 8\beta}{16\beta} \equiv \frac{(s - 8\beta)^2 - 1 + 8\beta}{16\beta} + 1 \pmod{2}.$$

We also ignore values of $b$ which are four or eight times an odd number, as then the congruence has no solutions, for, $s$ being odd,

$$s^2 \equiv 1 \pmod{8}.$$

Finally, when $b$ is twice an odd number, say $2\beta$, $s$ must be even and then

$$\frac{s^2 - 1 + \frac{1}{2}b}{b} \equiv \frac{\beta - 1}{2} \pmod{2}.$$

Since a unique correspondence exists between the solutions of the congruences

$$s^2 \equiv 1 - \beta \pmod{2\beta}, \qquad s^2 \equiv 1 \pmod{\beta},$$

we have

$$B_1 = \frac{1}{2^{\frac{1}{2}(r-1)}} \sum_{b \text{ odd}} \frac{(-1)^{\frac{1}{2}(b-1)} f(b)}{b^{\frac{1}{2}(r-1)}}.$$

Hence, from the formulae for $A_1$, it follows that the part of the coefficient of $q$ in $\chi$ arising from $B_1$ can be written as

$$\frac{\pi^{\frac{1}{2}r} B_1}{\Gamma(\frac{1}{2}r) S_{\frac{1}{2}(r-1)}} = \frac{\pi^{\frac{1}{2}r} \sum_{m \text{ odd}} (-1)^{\frac{1}{2}(m-1)} m^{-\frac{1}{2}(r-1)}}{2^{\frac{1}{2}(r-1)} \Gamma(\frac{1}{2}r) \sum_{m \text{ odd}} m^{-(r-1)}}.$$

Hence when $r \equiv 1 \pmod{4}$ the coefficient of $q$ in $\chi$ is

$$\frac{\pi^{\frac{1}{2}r}}{\Gamma(\frac{1}{2}r)} \left[1 + \frac{(-1)^{\frac{1}{4}(r-1)}}{2^{\frac{1}{2}(r-1)}} + \frac{2(-1)^{\frac{1}{4}(r-1)}}{4^{\frac{1}{2}(r-1)}}\right] \frac{\sum_{m \text{ odd}} m^{-\frac{1}{2}(r-1)}}{\sum_{m \text{ odd}} m^{-(r-1)}},$$

but when $r \equiv 3 \pmod{4}$ it is given by

$$\frac{\pi^{\frac{1}{2}r}}{\Gamma(\frac{1}{2}r)} \left[1 + \frac{1}{2^{\frac{1}{2}(r-1)}}\right] \frac{\sum_{m \text{ odd}} (-1)^{\frac{1}{2}(m-1)} m^{-\frac{1}{2}(r-1)}}{\sum_{m \text{ odd}} m^{-(r-1)}}.$$

Inserting the values of these series in the various cases*, we have

$$\theta_{00}{}^r = \chi + C_r \theta_{00}{}^{r-8} \theta_{01}{}^4 \theta_{10}{}^4,$$

where

$$C_9 = \tfrac{2}{17}, \quad C_{11} = \tfrac{22}{31}, \quad C_{13} = \tfrac{871}{691}, \quad C_{15} = \tfrac{861}{508}.$$

Let us examine this solution of the problem of finding the number of representations of a number as a sum of $r$ squares. In the case of eleven squares, say, we have

$$\theta_{00}{}^{11} = \chi + \tfrac{22}{31} \theta_{00}{}^3 \theta_{01}{}^4 \theta_{10}{}^4,$$

but the expansion of functions such as $\theta_{00}{}^3 \theta_{01}{}^4 \theta_{10}{}^4$ is in general a problem in itself. They however can be replaced by other functions, the coefficients of whose expansions as a power series in $q$ possess an arithmetical significance. The general method is given in a paper published

* See for example Chrystal, *Algebra*, vol. II. pp. 231, 342, 366.

several years ago*. For our present needs, we can start from the identity

$$\theta_{00}\left(\frac{\alpha x}{\omega}, -\frac{1}{\omega}\right)\theta_{00}\left(\frac{\beta x}{\omega}, -\frac{1}{\omega}\right)\theta_{00}\left(\frac{\gamma x}{\omega}, -\frac{1}{\omega}\right) = (\sqrt{-i\omega})^3 e^{\pi i x^2(\alpha^2+\beta^2+\gamma^2)/\omega}\,\theta_{00}(\alpha x, \omega)\,\theta_{00}(\beta x, \omega)\,\theta_{00}(\gamma x, \omega).$$

If $\alpha$, $\beta$ and $\gamma$ are constants satisfying the equation

$$\alpha^2+\beta^2+\gamma^2=0,$$

there is no difficulty in showing that an invariant for completing the solution in the case of eleven squares is given by the coefficient of $x^4$ on the right-hand side, omitting the factor $(\sqrt{-i\omega})^3$. This invariant, omitting an unimportant numerical factor, can be written as

$$\sum_{\pm\infty,}\sum_{\pm\infty,}\sum_{\pm\infty}(\alpha x+\beta y+\gamma z)^4 q^{x^2+y^2+z^2}.$$

By taking different values of $\alpha$, $\beta$ and $\gamma$, and combining these invariants linearly, we can deduce the more general invariant

$$\sum_{\pm\infty,}\sum_{\pm\infty,}\sum_{\pm\infty} f(x, y, z)\, q^{x^2+y^2+z^2},$$

where $f(x, y, z)$ is a polynomial solution, of degree four in $x$, $y$ and $z$, of

$$\frac{\partial^2 f}{\partial x^2}+\frac{\partial^2 f}{\partial y^2}+\frac{\partial^2 f}{\partial z^2}=0.$$

Take then $$f(x, y, z)=x^4-6x^2y^2+y^4.$$

It is easy to see that

$$\sum_{\pm\infty,}\sum_{\pm\infty,}\sum_{\pm\infty}(x^4+y^4-6x^2y^2)\, q^{x^2+y^2+z^2}=C\theta_{00}{}^3\theta_{01}{}^4\theta_{10}{}^4,$$

where $C$ is a constant, obviously $\frac{1}{4}$. We have now

$$\theta_{00}{}^{11}=\chi+\tfrac{88}{31}\sum_{\pm\infty,}\sum_{\pm\infty,}\sum_{\pm\infty}(x^4+y^4-6x^2y^2)\, q^{x^2+y^2+z^2}.$$

It is worthy of note that, as no numbers of the forms $8m+7$ or $4(8m+7)$ can be represented as a sum of three squares, the number of representations of such numbers as a sum of eleven squares can be expressed in a very simple form. In the first case for example, when $M=8m+7$ and has no squared factors, the number of representations is equal to

$$\frac{495\times 192}{31}\sum_{n \text{ odd}} M^{\frac{9}{2}}(-1)^{\frac{1}{2}(n-1)}\left(\frac{M}{n}\right)\frac{1}{n^5}.$$

Similarly, in the cases of thirteen and fifteen squares, we have

$$\theta_{00}{}^5\,\theta_{01}{}^4\,\theta_{10}{}^4=4\sum_{\pm\infty}\sum\sum\sum\sum(x^4+y^4-6x^2y^2)\, q^{x^2+y^2+z^2+t^2+u^2},$$

while $\theta_{00}{}^7\,\theta_{01}{}^4\,\theta_{10}{}^4$ is equal to the corresponding series with seven squares in the exponent of $q$.

In the case of nine squares, we can express $\theta_{00}\,\theta_{01}{}^4\,\theta_{10}{}^4$ by means of the representations of a number as a sum of seven squares, if we note that

$$\theta_{11}{}'=\pi\theta_{00}\,\theta_{01}\,\theta_{10}.$$

A simpler result might perhaps have been expected in this case.

* "Theta functions in the theory of the modular functions." *Quarterly Journal of Pure and Applied Mathematics*, vol. XLVI. p. 97.

# XVIII. *The Hydrodynamical Theory of the Lubrication of a Cylindrical Bearing under Variable Load, and of a Pivot Bearing.*

By W. J. HARRISON, M.A., Fellow of Clare College, Cambridge.

[*Received* 24 April 1919. *Read* 27 October 1919.]

Some of the results given in this paper were obtained in 1913, as stated in the author's previous paper on this subject*, but absence from Cambridge has delayed the completion of the work. It is thought that the results now presented add considerable interest to the theory of lubrication, as by their aid it is possible to picture what happens when either the speed or load is changed, or the bearing is subjected to small vibrations or impulses.

The good agreement shown in the previous paper to exist between theory and experiment in the case of an air-lubricated bearing may be accepted as establishing the trustworthiness of the general theory, so that the results now obtained may be regarded as reliable within limits, even if they can never be practically verified.

The second part of this paper has arisen from a paper on "Experiments on the Friction of a Pivot Bearing†."

*Cylindrical Bearing under Variable Load.*

Under variable load the relative position of the shaft and bearing must change. Let the shaft rotate with angular velocity $U/a$, where $a$ is its radius, and let it have a linear velocity $V$ relative to the bearing in a direction making an angle $\alpha$ with the line of centres $OO'$ of the two surfaces as shown in Fig. 1.

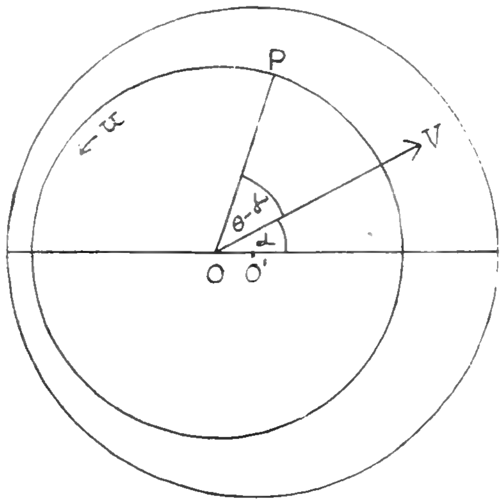



Fig. 1.

* *Transactions Cambridge Philosophical Society*, Vol. XXII., 1913, pp. 39—54.

† *Proc. Inst. Mech. Eng.*, March, 1891, The Fourth Report of the Research Committee on Friction.

If reference be made to the previous paper it will be found that equations (1), (2), (3), (5) (6) remain unchanged. The boundary conditions (4) become

$$u = U, \quad v = V\cos(\theta - \alpha), \quad \text{when } y = 0,$$
$$u = 0, \quad v = 0, \qquad \text{when } y = h.$$

We have

$$u = \frac{1}{2\mu}\frac{\partial p}{\partial x}y(y-h) + U\frac{h-y}{h},$$

and

$$\int_0^h \frac{\partial u}{\partial x}dy = -\Big[v\Big]_0^h = V\cos(\theta - \alpha).$$

Therefore

$$\frac{\partial}{\partial x}\left(h^3\frac{\partial p}{\partial x}\right) = 6\mu\left[U\frac{\partial h}{\partial x} - 2V\cos(\theta - \alpha)\right],$$

hence

$$h^3\frac{\partial p}{\partial x} = 6\mu Uh - 12\mu Va\sin(\theta - \alpha) + k,$$

or

$$\frac{1}{a}\frac{\partial p}{\partial \theta} = \frac{6\mu U\eta(1 + c\cos\theta) - 12\mu Va\sin(\theta - \alpha) + k}{\eta^3(1 + c\cos\theta)^3},$$

where $k$ is a constant of integration.

Now $p$ is a single-valued function of $\theta$, and, therefore, on integration it will be found that

$$k = -12\mu(1 - c^2)\{U\eta + 2V(a/c)\sin\alpha\}/(2 + c^2) + 12\mu V(a/c)\sin\alpha.$$

Accordingly

$$p = \frac{6\mu a}{\eta^3(1 + c\cos\theta)^2}\left[\{U\eta + 2V(a/c)\sin\alpha\}\, c\sin\theta\,\frac{2 + c\cos\theta}{2 + c^2} - V(a/c)\cos\alpha\right] + C,$$

where $C$ is a constant of integration.

Due to the normal pressure the forces acting on the shaft per unit length are $R$ in a downward direction perpendicular to $OO'$, and $S$ along $O'O$, where

$$R = \frac{12\pi\mu a^2 c}{\eta^2(2 + c^2)(1 - c^2)^{\frac{1}{2}}}\left(U + \frac{2Va\sin\alpha}{c\eta}\right),$$
$$S = 12\pi\mu a^3 V\cos\alpha/\{\eta^3(1 - c^2)^{\frac{3}{2}}\}.$$

The tangential traction $f$, acting on the shaft in a sense opposite to that of the rotation, is given by

$$f = -\mu\left(\frac{\partial u}{\partial y} + \frac{\partial v}{\partial x}\right), \quad y = 0,$$
$$= \mu\left[\frac{U}{h} + \frac{h}{2\mu a}\frac{\partial p}{\partial \theta} + \frac{V\sin(\theta - \alpha)}{a}\right].$$

The term $V\sin(\theta - \alpha)/a$, is clearly negligible in comparison with the other two.

This traction will give a resultant force acting on the shaft, depending on $U/\eta$ and $V/\eta^2$, and this force can be neglected in comparison with $R$ and $S$ above, which depend on $U/\eta^2$ and $V/\eta^3$. In fact the part depending on $U/\eta$ actually vanishes, as stated on p. 43 of the previous paper.

The traction also gives a couple of moment $M$ per unit length of the shaft, where

$$M = 4\pi\mu a^2[(1 + 2c^2)U\eta + 3Vac\sin\alpha]/\{\eta^2(2 + c^2)(1 - c^2)^{\frac{1}{2}}\}.$$

The equations determining the relative motion of the shaft and bearing may now be written down on the justifiable assumption that, if the shaft be in any position relative to the bearing and load, the velocity $V$ is instantaneously adjusted so as to enable the shaft to carry the load.

Let $W$ be a constant vertical load, and $w$ a variable load making an angle $\phi$ with the horizontal at time $t$. Also at time $t$ let $OO'$ make an angle $\psi$ with the horizontal, as in Fig. 2.

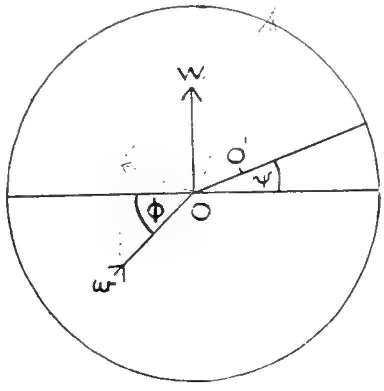



Fig. 2.

Then
$$\left.\begin{aligned} R &= W\cos\psi + w\sin(\phi-\psi) \\ S &= W\sin\psi + w\cos(\phi-\psi) \end{aligned}\right\} \qquad \text{(A)}.$$

Now
$$V\cos\alpha = -\eta\frac{dc}{dt},$$

and
$$V\sin\alpha = -\eta c\frac{d\psi}{dt}.$$

Let $c_1$ be the value of $c$ giving the position of the shaft in the case of steady motion with velocity $U$ and load $W$, so that
$$\frac{12\pi\mu a^2 c_1 U}{\eta^2(2+c_1^2)(1-c_1^2)^{\frac{1}{2}}} = W,$$

and
$$\psi = 0.$$

Also write $\tau = (W\eta^2/12\pi\mu a^3)\,t$.

The equations (A) become, after substituting for $R$ and $S$ the expressions found above,
$$\left.\begin{aligned} -\frac{d\psi}{d\tau} &= \{\cos\psi + (w/W)\sin(\phi-\psi)\}\,(1+\tfrac{1}{2}c^2)(1-c^2)^{\frac{1}{2}}/c - (1+\tfrac{1}{2}c_1^2)(1-c_1^2)^{\frac{1}{2}}/c_1 \\ -\frac{dc}{d\tau} &= (1-c^2)^{\frac{3}{2}}\{\sin\psi + (w/W)\cos(\phi-\psi)\} \end{aligned}\right\} \quad \text{(B)}.$$

PROBLEM 1. *The load is assumed to be constant, but owing to some cause, for example, a previous change of load, the position of the shaft is not the one proper for the load. It is required to trace the motion of the shaft relative to the bearing, if its angular velocity is maintained constant.*

The equations (B) determining the motion become
$$\left.\begin{aligned} -\frac{d\psi}{d\tau} &= \{(1+\tfrac{1}{2}c^2)(1-c^2)^{\frac{1}{2}}\cos\psi\}/c - (1+\tfrac{1}{2}c_1^2)(1-c_1^2)^{\frac{1}{2}}/c_1 \\ -\frac{dc}{d\tau} &= (1-c^2)^{\frac{3}{2}}\sin\psi \end{aligned}\right\} \qquad \text{(C)}.$$

Write
$$-\frac{d\psi}{d\tau}=f(c)\cos\psi+k,$$
$$-\frac{dc}{d\tau}=F(c)\sin\psi.$$

Then
$$F(c)\sin\psi\,d\psi-f(c)\cos\psi\,dc=k\,dc.$$

Let $\lambda$ be an integrating factor of the left-hand side of this equation. We have
$$d[-\lambda F(c)\cos\psi]=k\lambda\,dc,$$
where
$$\frac{d\lambda F(c)}{dc}=\lambda f(c),$$
or
$$\frac{d\lambda}{\lambda}=\frac{f(c)-F'(c)}{F(c)}dc$$
$$=\frac{2+7c^2}{2c(1-c^2)}dc.$$

Hence a value of $\lambda$ is $c(1-c^2)^{-\frac{9}{4}}$.

Therefore
$$-\lambda F(c)\cos\psi=k\int c(1-c^2)^{-\frac{9}{4}}dc+B,$$
or
$$-c(1-c^2)^{-\frac{3}{4}}\cos\psi=\tfrac{2}{5}k(1-c^2)^{-\frac{5}{4}}+B,$$
or
$$\cos\psi=-\frac{2k}{5c(1-c^2)^{\frac{1}{2}}}-\frac{B(1-c^2)^{\frac{3}{4}}}{c}\quad\text{.............................(D)},$$
where $B$ is a constant of integration, and
$$k=-(1+\tfrac{1}{2}c_1^2)(1-c_1^2)^{\frac{1}{2}}/c_1.$$

Equation (D) determines the path of the centre $O$ of the shaft relative to the centre $O'$ of the bearing. In Figs. 3–6 graphs are given showing the relation between $c$ and $\psi$ for various values

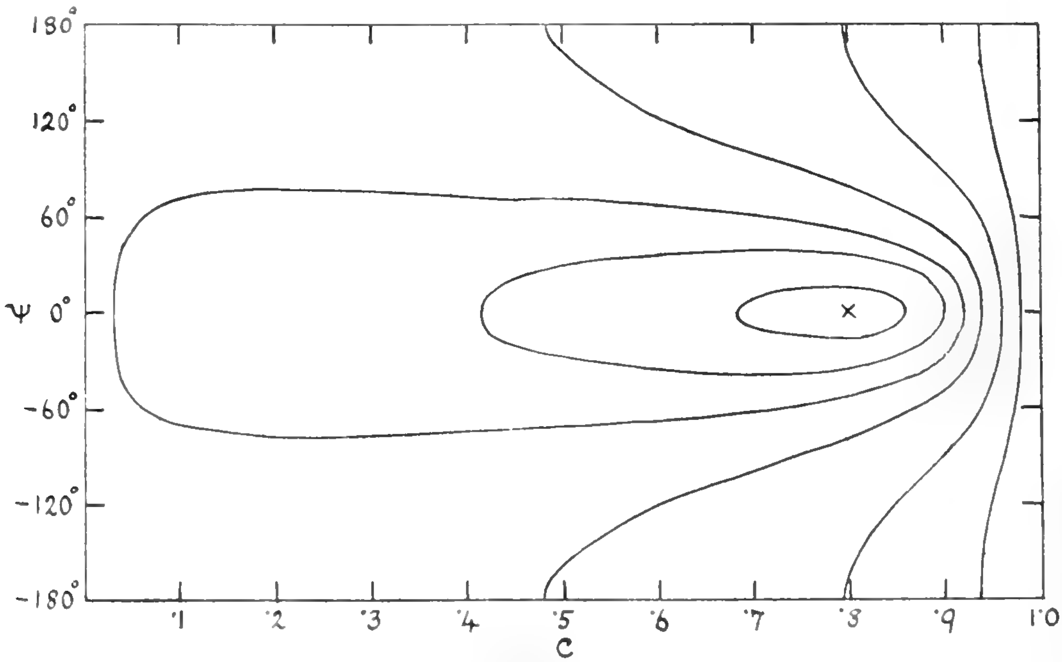



Fig. 3.

of $k$ and $B$. The value of $c_1$ assumed is shown by $X$ in the graph, and gives the position of the shaft proper for steady motion. The oval curves correspond to those cases in which $O$ describes a small closed curve not enclosing $O'$, and the other curves correspond to those cases in which the path of $O$ encloses $O'$.

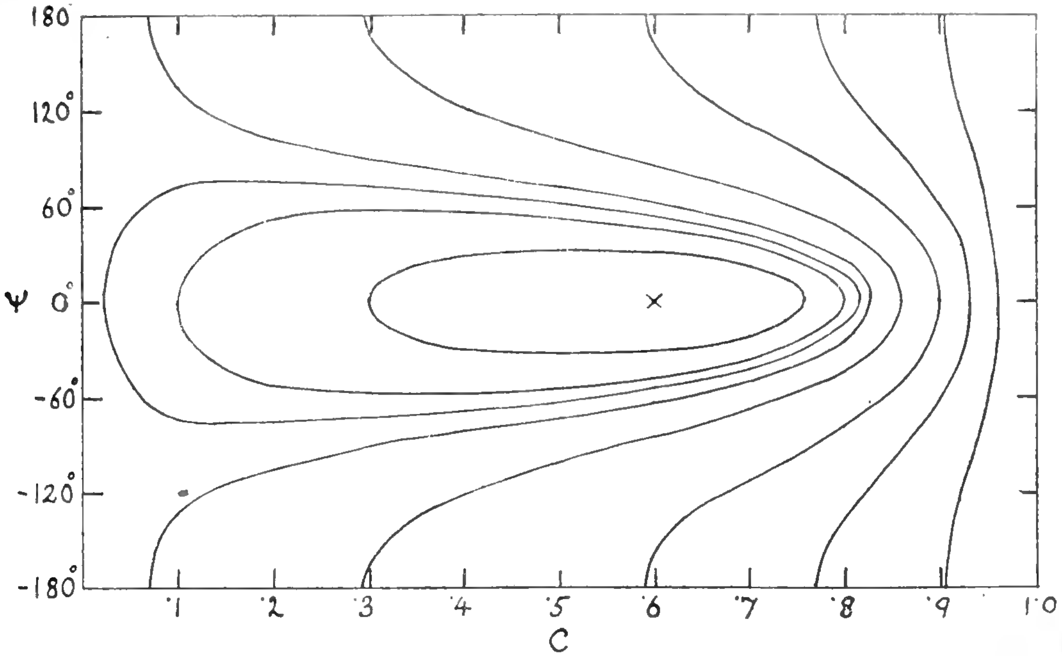



Fig. 4.

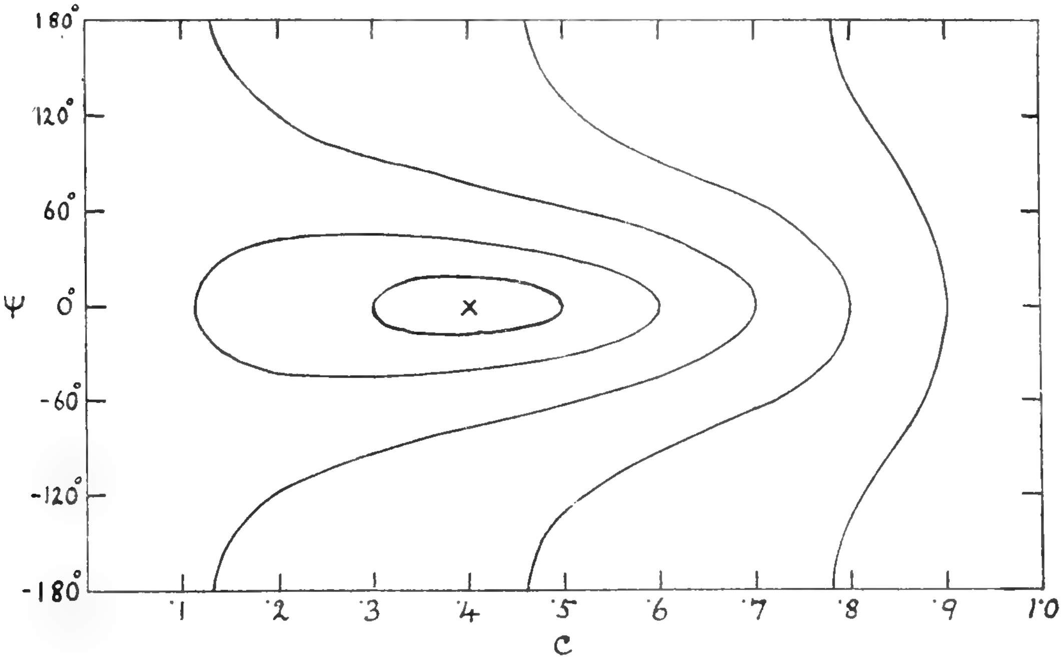



Fig. 5.

According to this analysis, if the shaft is displaced from its proper position it does not return to that position in the absence of further disturbances. This is no doubt incorrect, and is the result of a first approximation in which various factors have been neglected which would probably have the effect of restoring the steady motion after a time.

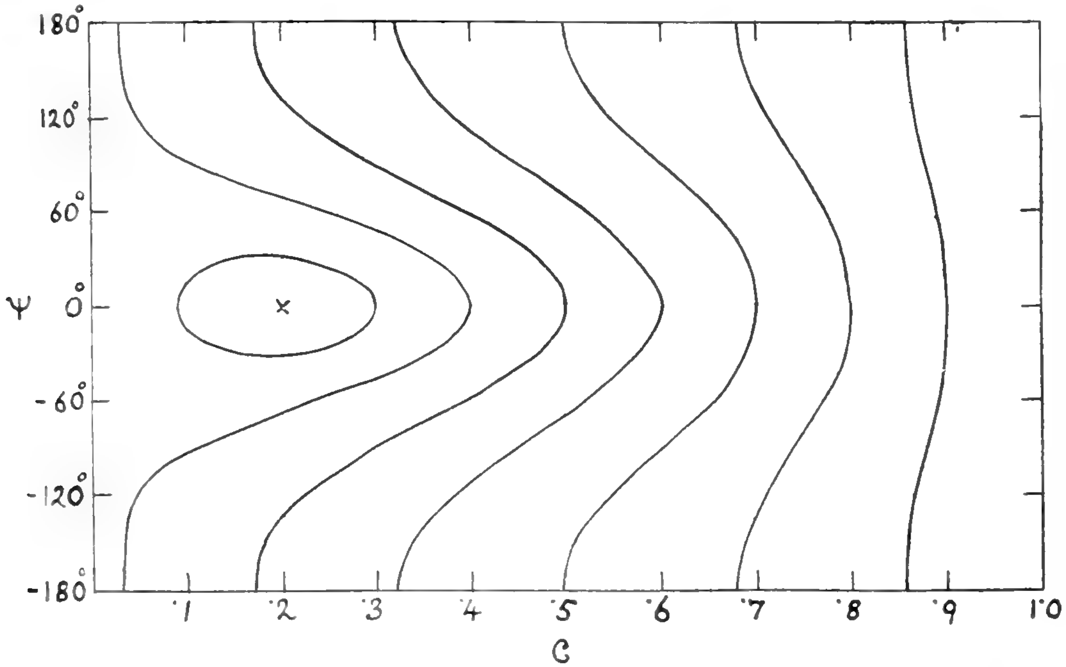



Fig. 6.

The graphs in Figs. 3–6 have been obtained by treating $c$ and $\psi$ as rectangular coordinates. If $c$ and $\psi$ be treated as polar coordinates the actual path of $O$ is obtained, as in Fig. 7, which corresponds to Fig. 5.

The data given in the following table may be used in connection with Figs. 3–6.

*Load per unit length carried by a shaft rotating steadily.*

| $c$ | ·1 | ·3 | ·5 | ·7 | ·9 |
|---|---|---|---|---|---|
| $\frac{W\eta^2}{12\pi\mu a^2 U}$ | ·050 | ·149 | ·257 | ·394 | ·735 |

Suppose, for example, that a shaft is carrying a load and rotating steadily in a position given by $c = \cdot 4$, $\psi = 0$, and the load is suddenly trebled, the proper value for $c$ is now ·8 approximately. Hence the motion of the shaft is approximately given by the intermediate oval curve in Fig. 3. It is assumed that the angular velocity is maintained constant.

The effect of varying the angular velocity of the shaft may be seen from the data already given, as such a change is equivalent to a variation in load as regards the relative position of the shaft and bearing. The effect of a gradual change of load or speed can be constructed roughly from the results given.

PROBLEM 2. *Given the shaft in any position, it is required to find the couple exerted on it by the traction due to the lubricant, the load being constant.*

The expression for the frictional couple $M$ has already been given.

$$M = 4\pi\mu a^2 \{(1+2c^2)U\eta + 3acV\sin\alpha\}/\{\eta^2(2+c^2)(1-c^2)^{\frac{1}{2}}\}$$

$$= 4\pi\mu a^2\left\{(1+2c^2)U - 3ac^2\frac{d\psi}{dt}\right\}\Big/\{\eta(2+c^2)(1-c^2)^{\frac{1}{2}}\}$$

$$= 4\pi\mu a^2\{(1+\tfrac{1}{2}c^2)U + c(2+c^2)(1-c^2)^{\frac{1}{2}}\eta^2 W\cos\psi/8\pi\mu a^2\}/\{\eta(2+c^2)(1-c^2)^{\frac{1}{2}}\}$$

$$= 2\pi\mu a^2 U[(1-c^2)^{-\frac{1}{2}} + 3cc_1\cos\psi/\{(1-c_1^2)^{\frac{1}{2}}(2+c_1^2)\}]\eta^{-1}.$$

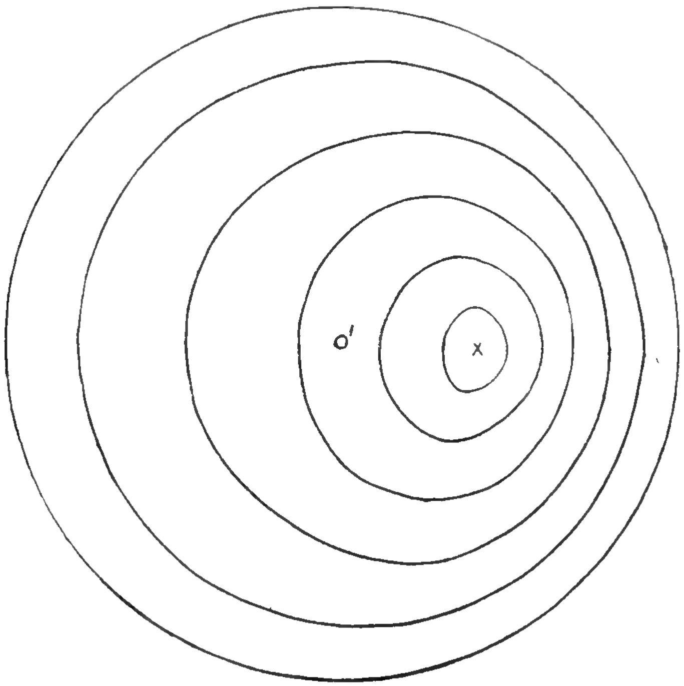



Fig. 7.

If $M_0$ is the value of the frictional couple when the shaft is rotating in its proper position for the given load and speed, then

$$M_0 = 4\pi\mu a^2(1+2c_1^2)U/\{\eta(2+c_1^2)(1-c_1^2)^{\frac{1}{2}}\}.$$

Hence

$$\frac{M}{M_0} = \frac{(1-c_1^2)^{\frac{1}{2}}(2+c_1^2) + 3cc_1(1-c^2)^{\frac{1}{2}}\cos\psi}{2(1+2c_1^2)(1-c^2)^{\frac{1}{2}}}.$$

The variation of $M/M_0$ with the position of the shaft can be seen from the following table:

$M/M_0$.

| $c_1$ | $c$ | $\psi$ = 0 | 60° | 90° | 120° | 180° |
|---|---|---|---|---|---|---|
| ·2 | ·2 | 1 | 0·97 | 0·94 | 0·91 | 0·88 |
| | ·5 | 1·20 | 1·13 | 1·07 | 1·00 | 0·93 |
| | ·8 | 1·76 | 1·65 | 1·54 | 1·43 | 1·32 |
| | ·95 | 3·23 | 3·10 | 2·97 | 2·83 | 2·70 |
| ·5 | ·2 | 0·77 | 0·72 | 0·67 | 0·62 | 0·57 |
| | ·5 | 1 | 0·88 | 0·75 | 0·63 | 0·50 |
| | ·8 | 1·49 | 1·29 | 1·09 | 0·88 | 0·68 |
| | ·95 | 2·55 | 2·32 | 2·08 | 1·84 | 1·61 |
| ·8 | ·2 | 0·46 | 0·41 | 0·36 | 0·30 | 0·25 |
| | ·5 | 0·66 | 0·53 | 0·40 | 0·27 | 0·14 |
| | ·8 | 1 | 0·79 | 0·58 | 0·37 | 0·16 |
| | ·95 | 1·62 | 1·37 | 1·12 | 0·86 | 0·61 |

The use of this table in conjunction with Figures 3–6 will indicate the extent to which the frictional couple may be varied by causes which result in a displacement of the shaft. For example, if a heavy load be greatly lightened, the resulting mean frictional couple will be considerably greater than the frictional couple in the state of steady motion proper to the load as lightened.

PROBLEM 3. *The load is constant, but the shaft is displaced slightly from its proper position for steady motion; to find the motion.*

These cases of motion would give rise to small oval curves in Figures 3–6, but they can be treated separately by a method of approximation.

In equations (D), write $c = c_1 + c'$, and treat $c'$ and $\psi$ as small. We find

$$\frac{d\psi}{d\tau} = \frac{1 - \frac{1}{2}c_1^2 + c_1^4}{c_1^2(1 - c_1^2)^{\frac{1}{2}}} c' = \kappa_1 c',$$

$$\frac{dc'}{d\tau} = -(1 - c_1^2)^{\frac{3}{2}} \psi = -\kappa_2 \psi.$$

Hence
$$\frac{d^2c'}{d\tau^2} = -\kappa_1\kappa_2 c'.$$

Therefore
$$c' = \alpha \cos\{(\kappa_1\kappa_2)^{\frac{1}{2}}\tau + \epsilon\},$$

$$\psi = -\alpha(\kappa_1/\kappa_2)^{\frac{1}{2}} \sin\{(\kappa_1\kappa_2)^{\frac{1}{2}}\tau + \epsilon\},$$

where
$$\kappa_1\kappa_2 = (1 - \tfrac{1}{2}c_1^2 + c_1^4)(1 - c_1^2)/c_1^2,$$

$$\kappa_1/\kappa_2 = (1 - \tfrac{1}{2}c_1^2 + c_1^4)/\{c_1^2(1 - c_1^2)^2\}.$$

The motion of $O$ is periodic, and the period can be found by transforming back from $\tau$ to $t$.

$$\begin{aligned}\text{Periodic time} &= \{2\pi/(\kappa_1\kappa_2)^{\frac{1}{2}}\}\, 12\pi\mu a^3/(W\eta^2)\\ &= \{2\pi/(\kappa_1\kappa_2)^{\frac{1}{2}}\}\, a\,(2+c_1^2)(1-c_1^2)^{\frac{1}{2}}/(c_1 U)\\ &= \frac{60\,(2+c_1^2)}{n\,(1-\frac{1}{2}c_1^2+c_1^4)},\end{aligned}$$

where $n$ is number of revolutions per minute of the shaft.

| $c$ | ·1 | ·3 | ·5 | ·7 | ·9 |
|---|---|---|---|---|---|
| $\frac{60\,(2+c_1^2)}{1-\frac{1}{2}c_1^2+c_1^4}$ | 121 | 130 | 145 | 150 | 135 |

PROBLEM 4. *To find the effect of a small additional oscillating load.*

In equations (B) write $c = c_1 + c'$, where $c_1$ gives the steady position under the constant load only, and assume $c'$ and $\psi$ to be small.

We have

$$\left.\begin{aligned}\frac{d\psi}{d\tau} &= \kappa_1 c' - \kappa_3 (w/W)\sin\phi\\ \frac{dc'}{d\tau} &= -\kappa_2\psi - \kappa_4(w/W)\cos\phi\end{aligned}\right\} \quad\ldots\ldots\ldots\ldots\ldots\ldots (\text{E}),$$

where

$$\begin{aligned}\kappa_1 &= (1-\tfrac{1}{2}c_1^2+c_1^4)/\{c_1^2(1-c_1^2)^{\frac{1}{2}}\},\\ \kappa_3 &= (1+\tfrac{1}{2}c_1^2)(1-c_1^2)^{\frac{1}{2}}/c_1,\\ \kappa_2 &= \kappa_4 = (1-c_1^2)^{\frac{3}{2}}.\end{aligned}$$

If we assume $w/W \propto \cos p\tau$, and that $\phi$ is constant, equations (E) fall under a well-known type, and the solution is of the form

$$c' = \alpha\cos\{(\kappa_1\kappa_2)^{\frac{1}{2}}\tau+\epsilon\} + \beta\cos(p\tau+\epsilon'), \text{ etc.}$$

Thus the effect can be found of small periodic variations in speed or load, or of small periodic impulses.

*Theory of Lubrication of a Pivot Bearing.*

A Pivot Bearing consists of a vertical cylindrical shaft, capable of rotation about its axis, having a plane end which bears on a horizontal plane surface. When the bearing is lubricated the two surfaces will be separated by a film of oil, and maintained at a uniform distance apart.

The case shown in Fig. 8 will be considered first.

The bearing has a central channel of radius $r_1$ for the purpose of supplying oil. The radius of the shaft is $r_2$. It will be shown that efficient lubrication is not possible under these simple conditions, but a consideration of this case affords an introduction to the more extended theory given later.

Let $(x, y)$ be coordinates in the plane of the moving surface, $z$ the coordinate at right angles to the plane drawn towards the fixed surface. Let $(u, v, w)$ be component velocities at any point in the liquid film.

If $h$ is the distance between the two surfaces, the boundary conditions are of the form

$$u = U, \quad v = V, \quad w = 0, \quad \text{when } z = 0,$$
$$u = 0, \quad v = 0, \quad w = 0, \quad \text{when } z = h.$$

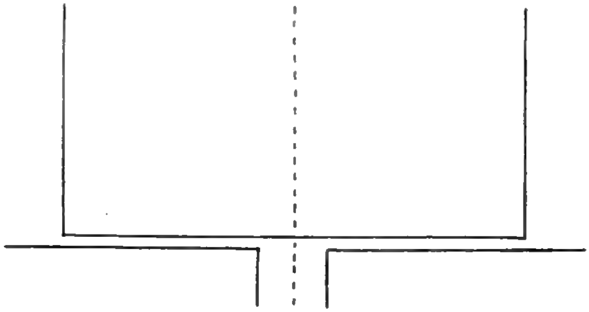


Fig. 8.

With the usual approximations

$$\frac{\partial p}{\partial x} = \mu \frac{\partial^2 u}{\partial z^2},$$
$$\frac{\partial p}{\partial y} = \mu \frac{\partial^2 v}{\partial z^2},$$
$$\frac{\partial p}{\partial z} = 0.$$

Hence
$$u = \frac{1}{2\mu} \frac{\partial p}{\partial x} z (z - h) + U \frac{h - z}{h},$$
$$v = \frac{1}{2\mu} \frac{\partial p}{\partial y} z (z - h) + V \frac{h - z}{h}.$$

Now
$$\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} + \frac{\partial w}{\partial z} = 0,$$

and therefore
$$\int_0^h \left\{ \frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} \right\} dz = -\Big[ w \Big]_0^h = 0.$$

Hence
$$\frac{\partial}{\partial x}\left(h^3 \frac{\partial p}{\partial x}\right) + \frac{\partial}{\partial y}\left(h^3 \frac{\partial p}{\partial y}\right) = 6\mu \left(\frac{\partial h U}{\partial x} + \frac{\partial h V}{\partial y}\right),$$

and in the present problem $U = -\omega y$, $V = \omega x$, hence

$$\frac{\partial^2 p}{\partial x^2} + \frac{\partial^2 p}{\partial y^2} = 0.$$

Taking the origin at the centre of the plane end of the shaft, and using cylindrical coordinates $(r, \theta, z)$, we have

$$\frac{\partial^2 p}{\partial r^2} + \frac{1}{r}\frac{\partial p}{\partial r} + \frac{1}{r^2}\frac{\partial^2 p}{\partial \theta^2} = 0,$$
$$u = \frac{1}{2\mu}\frac{\partial p}{\partial r} z (z - h).$$
$$v = \frac{1}{2\mu r}\frac{\partial p}{\partial \theta} z (z - h) + \omega r \frac{h - z}{h},$$

where $\omega$ is the angular velocity of the shaft.

In the case at present under consideration $u$, $v$, $p$ are functions of $r$ only, so that

$$p = A + B \log r,$$
$$u = Bz(z-h)/(2\mu r),$$
$$v = \omega r(h-z)/h.$$

The pressure satisfies the conditions

$$p = p_1 \text{ at } r = r_1,$$
$$p = p_2 \text{ at } r = r_2.$$

Hence
$$A = (p_1 \log r_2 - p_2 \log r_1)/\{\log(r_2/r_1)\},$$
$$B = (p_2 - p_1)/\{\log(r_2/r_1)\}.$$

Thus the pressure $p$, and consequently the weight supported, are independent of the distance between the two surfaces. The flow of oil, if it takes place at all, is dependent on the maintenance of a difference of pressure between the boundaries $r=r_1$ and $r=r_2$. Under these circumstances effective lubrication is not possible. This difficulty can be surmounted by the introduction of radial grooves cut in the bearing surface to serve as oil channels. Such a modification introduces considerable complexity, and it will be an advantage to consider the case of two parallel plane surfaces of which one has a motion of translation, but no rotation, relative to the other.

In order to ensure effective lubrication it is necessary to have oil channels cut in the bearing at right angles to the direction of motion, these channels having one or both edges chamfered off. In the case of reciprocating motion it is desirable to have both edges chamfered, but for the present purpose it is sufficient to consider the case in which only one edge is so treated, as in Fig. 9. The object of considering this problem first is to obtain suitable pressure conditions at the edges of a groove. There is a further interest in a consideration of this case in that, hitherto, the theory of the lubrication of plane surfaces has only been presented on the supposition that they are inclined to each other.

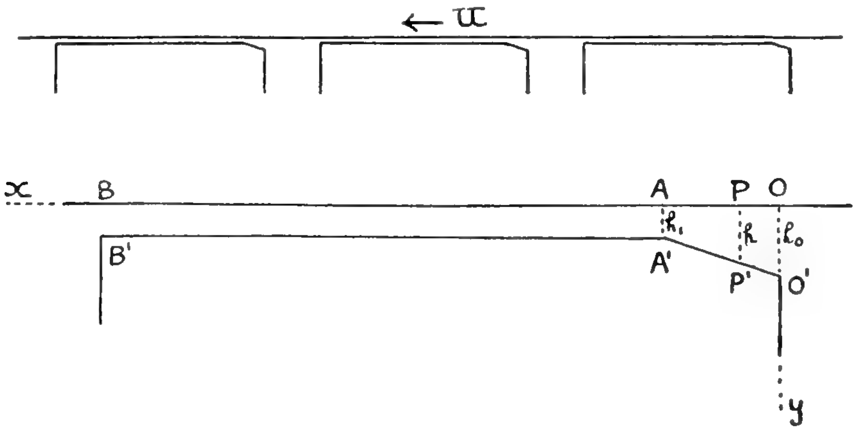



Fig. 9.

Let $OA = a$, $AB = b$, and $h$ the distance between the surfaces.

For
$$0 < x < a, \qquad h = h_0 - (h_0 - h_1)\, x/a,$$
$$a < x < a+b, \qquad h = h_1.$$

Let $p$ be the pressure and $u$ the velocity at any point in the liquid film, $\Pi$ the pressure in an oil channel. The motion is given by

$$h^3 \frac{\partial p}{\partial x} = 6\mu U h + C \quad \text{.......................................(1)},$$

$$u = \frac{1}{2\mu}\frac{\partial p}{\partial x} y(y-h) + U\frac{h-y}{h} \quad \text{..............................(2)},$$

where $C$ is a constant.

From (1)

$$p = \left[\frac{6\mu U}{h} + \frac{C}{2h^2}\right]\frac{a}{h_0-h_1} + B, \quad 0<x<a \quad \text{..........................(3)},$$

$$p = \left[\frac{6\mu U}{h_1^2} + \frac{C}{h_1^3}\right](x-a) + A, \quad a<x<a+b \quad \text{....................(4)},$$

where $A$ and $B$ are constants of integration.

The conditions are:

(a) $p$ is continuous, $x=a$, giving

$$A = \left[\frac{6\mu U}{h_1} + \frac{C}{2h_1^2}\right]\frac{a}{h_0-h_1} + B \quad \text{..............................(5)},$$

(b) $p=\Pi$, $x=a+b$, giving

$$\Pi = \left[\frac{6\mu U}{h_1^2} + \frac{C}{h_1^3}\right]b + A \quad \text{.......................................(6)},$$

(c) $p=\Pi$, $x=0$, giving

$$\Pi = \left[\frac{6\mu U}{h_0} + \frac{C}{2h_0^2}\right]\frac{a}{h_0-h_1} + B \quad \text{..............................(7)}.$$

Now, although $h_0-h_1$ is small, $h_1$ will in general be small compared with $h_0$ or $h_0-h_1$, and $a$ small compared with $b$. The following approximate values of the constants $A, B, C$ have been calculated on this supposition, and will be satisfactory in general, although in exceptional circumstances, such as a case of a very light load, it might be necessary to include other terms. We have

$$A = \Pi + 3\mu U a(h_0-h_1)/(h_0^2 h_1),$$
$$B = \Pi - 6\mu U a(h_0 + \tfrac{1}{2}h_1)/h_0^3,$$
$$C = -6\mu U h_1(bh_0^2 + \tfrac{1}{2}ah_1h_0 - \tfrac{1}{2}ah_1^2)/(bh_0^2).$$

Hence, with sufficient approximation

$0<x<a$,

$$p = \Pi + \frac{6\mu U a}{h_0}\left[\left(1+\frac{h_1}{h_0}\right)\left(\frac{1}{h} - \frac{1}{2}\frac{h_1}{h^2}\right) - \frac{1}{h_0}\left(1+\frac{1}{2}\frac{h_1}{h_0}\right)\right],$$

$$\frac{\partial p}{\partial x} = 6\mu U\left(\frac{1}{h^2} - \frac{h_1}{h^3}\right),$$

$$u = \frac{1}{2\mu}\frac{\partial p}{\partial x} y(y-h) + U\frac{h-y}{h}.$$

$a<x<a+b$,

$$p = \Pi + 3\mu U a(h_0-h_1)[1-(x-a)/b]/(h_0^2h_1),$$

$$\frac{\partial p}{\partial x} = -3\mu U a(h_0-h_1)/(bh_0^2h_1),$$

$$u = \frac{1}{2\mu}\frac{\partial p}{\partial x} y(y-h_1) + U\frac{h_1-y}{h_1}.$$

The weight supported per unit breadth by the portion of the surface between two grooves is

$$\frac{6\mu U ab}{4h_1h_0}\left[1 - \frac{h_1}{h_0} + \frac{ah_1}{bh_0}\left(1+2\frac{h_1}{h_0}\right)\log_e\frac{h_0}{h_1} - \frac{ah_1}{2bh_0}\left(3+2\frac{h_1}{h_0}\right)\right].$$

The traction on the moving surface between two grooves per unit breadth is

$$\frac{\mu U b}{h_1}\left[1+\frac{4ah_1}{bh_0}\left(1+\frac{h_1}{h_0}\right)\log_e\frac{h_0}{h_1}-\frac{3ah_1}{2bh_0}\left(3-\frac{h_1}{h_0}\right)\right].$$

If $\frac{h_1}{h_0}$ is very small, as will usually be the case, we have as a first approximation

Weight supported $= 3\mu Uab/(2h_1h_0)$,
Traction $= \mu Ub/h_1$,
Coefficient of friction $= 2h_0/3a$.

Hence the coefficient of friction depends only on the ratio $h_0/a$ or $(h_0-h_1)/a$, provided $h_0/a$ is small, and that $h_1/h_0$ is small.

Returning to the case of a pivot bearing with radial grooves having chamfered edges, it may be assumed that $U=\omega r$, and that $a$ and $(h_0-h_1)$ vary as $r$. In the problem just considered the pressure at $BB'$ due to the chamfered edge is $3\mu Ua(h_0-h_1)/(h_0^2h_1)$, so that $p=\Pi+\frac{\mu k\omega r}{h}$ is a reasonable assumption for the pressure at the edge of a groove, where $h$ is the distance between the two surfaces and $k$ is a constant, but, by suitably choosing $a$ and $(h_0-h_1)$ as functions of $r$, $p$ may be made to take any required form.

Let there be four radial grooves along two diameters at right angles. In each sector measure $\theta$ from a bounding groove in the direction of rotation of the shaft.

Now
$$\frac{\partial^2 p}{\partial r^2}+\frac{1}{r}\frac{\partial p}{\partial r}+\frac{1}{r^2}\frac{\partial^2 p}{\partial\theta^2}=0,$$
hence we may assume

$$p=A+B\log r+C\theta+\Sigma\,(A_n r^n+B_n r^{-n})\cos n\theta+\Sigma\,(C_n r^n+D_n r^{-n})\sin n\theta,$$

with the boundary conditions

$$p=\Pi+\mu k\omega r/h,\quad \theta=0,\quad 0<r<a,$$
$$p=\Pi,\quad \theta=\pi/2,\quad 0<r<a,$$
$$p=\Pi,\quad r=a,\quad 0<\theta<\pi/2.$$

The appropriate solution is

$$p=\Pi+\frac{\mu k\omega}{h}r\cos\theta-\frac{2\mu k\omega a}{\pi h}\Sigma\frac{4n}{4n^2-1}\frac{r^{2n}}{a^{2n}}\sin 2n\theta,$$

where $n$ is integral.

The distribution of pressure over the sector is shown in the following table, and curves of equal pressure are shown in Fig. 10, drawn for the cases $(p-\Pi)\,h/(\mu k\omega a)=$ ·6, ·4, ·2, ·05.

$$\frac{(p-\Pi)h}{\mu k\omega a}.$$

| $r$ | $\theta$ 10° | 20° | 30° | 40° | 50° | 60° | 70° | 80° |
|---|---|---|---|---|---|---|---|---|
| ·8$a$ | ·41 | ·22 | ·13 | ·08 | ·05 | ·04 | ·02 | ·01 |
| ·6$a$ | ·45 | ·31 | ·22 | ·15 | ·11 | ·07 | ·04 | ·02 |
| ·4$a$ | ·34 | ·28 | ·22 | ·17 | ·13 | ·09 | ·06 | ·03 |
| ·2$a$ | ·19 | ·17 | ·15 | ·13 | ·10 | ·08 | ·05 | ·03 |

The weight supported

$$= 4\int_0^{\frac{1}{2}\pi}\int_0^a \frac{\mu k\omega}{h}\left[r\cos\theta - \frac{2a}{\pi}\Sigma\frac{4n}{4n^2-1}\frac{r^{2n}}{a^{2n}}\sin 2n\theta\right] r\,dr\,d\theta$$

$$= \frac{4k\omega\mu a^3}{h}\left(\frac{1}{3} - \frac{4}{\pi}\Sigma\frac{1}{(4n^2-1)(n+1)}\right), \text{ where } n \text{ is odd}$$

$$= \cdot 440\,(k\omega\mu a^3/h).$$

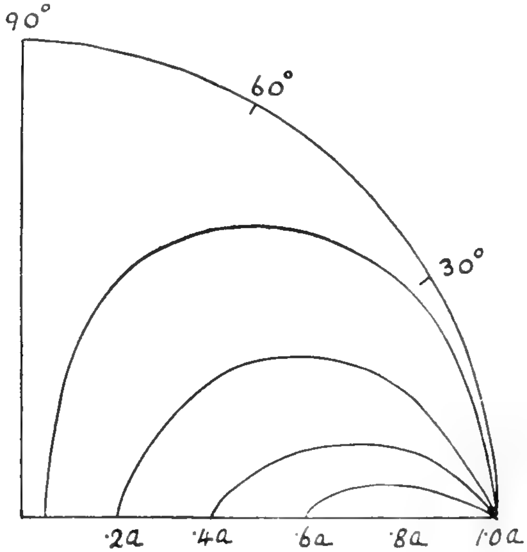



Fig. 10.

The frictional couple on the shaft per unit length

$$= 4\int_0^{\frac{1}{2}\pi}\int_0^a -\mu\left(\frac{\partial v}{\partial z}\right)_{z=0} r^2\,dr\,d\theta$$

$$= 4\int_0^{\frac{1}{2}\pi}\int_0^a \mu\left(\frac{\omega r}{h} + \frac{h}{2\mu r}\frac{\partial p}{\partial\theta}\right) r^2\,dr\,d\theta$$

$$= \tfrac{1}{2}\pi\mu\omega a^4/h - 4\int_0^a \tfrac{1}{2}\mu k\omega r^2\,dr$$

$$= \tfrac{1}{2}\pi\mu\omega a^4/h - \tfrac{2}{3}\mu k\omega a^3.$$

The term $\frac{2}{3}\mu k\omega a^3$ is of a smaller order than the other, and in fact terms of this order would occur if account were taken of that part of the frictional couple due to the chamfered edge.

$$\frac{\text{Frictional couple}}{\text{Weight supported}} = \tfrac{1}{2}\pi a/(\cdot 440k)$$

$$= 3\cdot 57\,(a/k).$$

Hence to a first order of approximation the ratio of the frictional couple to the weight carried is constant for all loads. The value of $k$ may be taken as $3/\epsilon$, where $\epsilon$ is the small angle measured in radians which the plane of chamfering of the edge of the groove makes with the plane of the bearing.

If the boundary $r=a$ is closed so that no flow of liquid takes place over it, the boundary conditions are

$$p=\Pi+\mu k\omega r/h, \quad \theta=0, \quad 0<r<a,$$
$$p=\Pi, \quad \theta=\tfrac{1}{2}\pi, \quad 0<r<a,$$
$$\frac{\partial p}{\partial r}=0, \quad r=a, \quad 0<\theta<\tfrac{1}{2}\pi.$$

The appropriate solution is

$$p=\Pi+\frac{\mu k\omega}{h} r\cos\theta-\frac{4\mu k\omega a}{\pi h}\Sigma\frac{r^{2n}}{(4n^2-1)\,a^{2n}}\sin 2n\theta,$$

where $n$ is a positive integer.

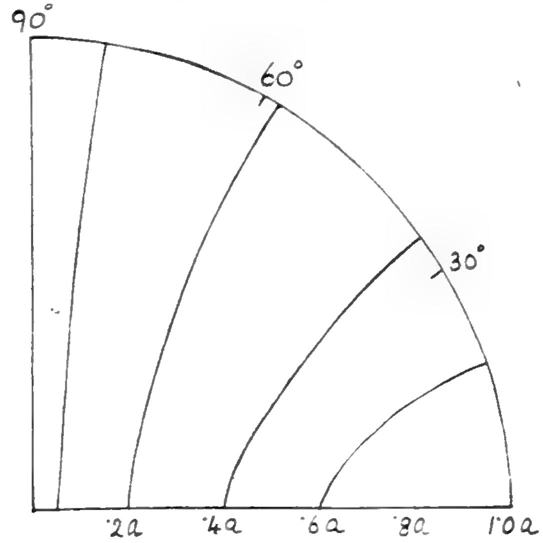



Fig. 11.

The distribution of pressure over the sector is shown in the following table, and curves of equal pressure are shown in Fig. 11, drawn for the cases

$$(p-\Pi)\,h/(\mu k\omega a)=·6,\ ·4,\ ·2,\ ·05.$$

| | $\frac{(p-\Pi)h}{\mu k\omega a}$ | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| $r$ | $\theta$ | | | | | | | |
| | 10° | 20° | 30° | 40° | 50° | 60° | 70° | 80° |
| ·8a | ·66 | ·53 | ·42 | ·34 | ·26 | ·19 | ·13 | ·06 |
| ·6a | ·53 | ·45 | ·38 | ·30 | ·24 | ·18 | ·12 | ·06 |
| ·4a | ·37 | ·33 | ·28 | ·24 | ·19 | ·14 | ·09 | ·05 |
| ·2a | ·19 | ·18 | ·16 | ·14 | ·12 | ·09 | ·06 | ·03 |

The weight supported

$$= \frac{4\mu k \omega a^3}{h} \left[ \frac{1}{3} - \frac{2}{\pi} \Sigma \frac{1}{n(n+1)(4n^2-1)} \right] \text{ (where } n \text{ is odd)}$$

$$= \cdot 902 \, (\mu k \omega a^3/h).$$

The frictional couple is the same as in the previous case. Hence

$$\frac{\text{Frictional couple}}{\text{Weight supported}} = \tfrac{1}{2} \pi a/(\cdot 902k)$$

$$= 1\cdot 74 \, (a/k).$$

A number of other cases can be solved. With regard to the experimental data given in the paper quoted from the *Proc. Inst. Mech. Eng.*, it would not appear possible to explain the fact that the flow of oil was stationary or even increased when the load was increased, the speed remaining unchanged.

Lord Rayleigh* has suggested a form for a footstep or pivot bearing, but without any theoretical investigation.

* *Phil. Mag.* (6), xxxv. p. 1, 1918.

## XIX. *On integers which satisfy the equation* $t^3 \pm x^3 \pm y^3 \pm z^3 = 0$.

By H. W. Richmond, M.A., King's College.

[*Received and read* 9 February 1920.]

### *Historical Sketch.* §§ 1—4.

1. The classical solution of this Diophantine problem*, i.e. of the equation

$$t^3 + x^3 + y^3 + z^3 = 0, \quad \text{..........................................(1)}$$

in integers positive or negative, is that given by Euler in 1754, viz.

$$\left.\begin{aligned} nt &= (3pr + 3qr - ps + 3qs)(s^2 + 3r^2) - (p^2 + 3q^2)^2, \\ nx &= (3pr - 3qr + ps + 3qs)(s^2 + 3r^2) + (p^2 + 3q^2)^2, \\ ny &= -(3pr + 3qr - ps + 3qs)(p^2 + 3q^2) + (s^2 + 3r^2)^2, \\ nz &= -(3pr - 3qr + ps + 3qs)(p^2 + 3q^2) - (s^2 + 3r^2)^2. \end{aligned}\right\} \quad \text{..................(A)}$$

Here the multiplier $n$ is merely equivalent to a symbol of proportionality. The unknowns $t, x, y, z$ are shewn to be proportional to homogeneous polynomials of the fourth degree in four parameters $p, q, r, s$. We need only consider solutions of the given equation consisting of sets of integers $t, x, y, z$ which have no factor common to all. Any such set can be obtained by giving to $p, q, r, s$ integer values positive or negative (having no common factor), calculating the numerical values of the functions on the right of the above equations and putting $n$ equal to their greatest common measure. An account of this and other methods of solution, due to Euler and earlier writers, is given in Sir T. L. Heath's *Diophantus of Alexandria*, 2nd edition (1910), pp. 101—102 and 329—334. Of other references it is here sufficient to mention (i) Fermat, *Oeuvres*, Vol. III, pp. 420 and 535, where 18 particular solutions, discovered by Frénicle, in which two of the numbers are positive and two negative, are quoted. (ii) Euler, *Commentationes Arithmeticae* (Petrograd, 1849), Vol. I, pp. 193—209, where further particular solutions will be found. Euler's statement of the problem is "To find sets of three cubes whose sum is a cube"; but he includes cases in which one of the three is negative when his formulae lead to such a result. This I have regarded as an admission that the two problems need not be separated, a view which seems in accordance with modern notions. I have therefore written the equation in the symmetrical form

$$t^3 + x^3 + y^3 + z^3 = 0, \quad \text{..........................................(1)}$$

so that all the solutions contain some positive and some negative numbers. In consequence, some changes of sign have been made in Euler's results.

* There is little or no doubt, in spite of a hiatus in the text, that Diophantus stated the theorem:—The difference of any two cubes is also the sum of two cubes; but his method of proof is lost. The numbers would as a rule be rational, not integral.

2. Euler's solution was modified in 1841 by Binet, *Comptes Rendus* (1841), Vol. XII, p. 248, who recognized that it was unnecessary to make use of as many as four parameters. By putting $r=0$ and $s=1$ in Euler's equations, he obtained a simpler but quite general solution,

$$\left.\begin{aligned} nt &= (3q-p)-(3q^2+p^2)^2, \\ nx &= (3q+p)+(3q^2+p^2)^2, \\ ny &= -(3q-p)(3q^2+p^2)+1, \\ nz &= -(3q+p)(3q^2+p^2)-1, \end{aligned}\right\} \quad \text{.................................(B)}$$

the parameters $p$ and $q$ being now allowed to assume any *rational* values positive or negative. This solution, known as the Euler-Binet solution, has come to be regarded as the standard solution of the Diophantine equation. It is obtained by another method in R. D. Carmichael's book on *Diophantine Analysis* (1915), pp. 62—66. Two other non-homogeneous solutions in terms of two parameters have recently been suggested by Schwering and Kühne respectively in the *Arch. Math. Phys.*, Series 3, Vols. II (1902) and III (1904); but they are shewn by Fujiwara in Vol. XIX of the same periodical (1912), and again in Vol. I of the *Tokohu Mathematical Journal*, to be at bottom equivalent to the Euler-Binet solution, and in some respects inferior to it.

3. Euler's solution can be obtained directly from the equations

$$\left.\begin{aligned} \frac{\omega t+\omega^2 x}{\omega z+\omega^2 y} &= \frac{\omega(p+q)+\omega^2(p-q)}{\omega(s-r)+\omega^2(s+r)}, \\ \frac{\omega^2 t+\omega x}{\omega^2 z+\omega y} &= \frac{\omega^2(p+q)+\omega(p-q)}{\omega^2(s-r)+\omega(s+r)}, \\ \frac{t+x}{z+y} &= -\frac{s^2+3r^2}{p^2+3q^2}, \end{aligned}\right\} \quad \text{..............................(C)}$$

where $\omega$ is one of the complex cube roots of unity. These equations imply that $t, x, y, z$ satisfy the given equation (1), as is seen at once by multiplication so as to eliminate $p, q, r, s$. On clearing them of fractions we have three linear equations in $t, x, y, z$, which can be solved, and will be found to lead to Euler's values of the ratios of $t, x, y, z$ to one another. To any chosen set of values of $p, q, r, s$ therefore corresponds a solution of the Diophantine equation. Also, when a set of values $t, x, y, z$ has been found which satisfy equation (1), the first two of equations (C) shew that an infinite number of corresponding sets of values of $p, q, r, s$ can be found which lead to that solution. If however we reduce the equations to the Euler-Binet form by putting $r=0$ and $s=1$, we have, after a little simplification,

$$\left.\begin{aligned} p+q &= (xy+tz-xz)/(y^2-yz+z^2), \\ p-q &= (xy+tz-ty)/(y^2-yz+z^2), \end{aligned}\right\} \quad \text{..............................(D)}$$

a result which proves that in the Euler-Binet formulae a unique pair of values can be found for $p$ and $q$ which will lead to any chosen values of $t, x, y, z$ that satisfy the given equation (1). Further, it will be seen that rational values of $p$ and $q$ correspond to the integer values of $t, x, y, z$ that we are seeking; and conversely.

4. Hermite in 1875 contributed a short note to the *Nouvelles Annales de Mathématiques*, Series 2, Vol. XI, p. 5, in which he pointed out the meaning of the Euler-Binet solution in connection with the surface of the third order represented in homogeneous coordinates by equation (1). I shall consider the geometric interpretation of equations (B) in section 5. In this note

Hermite takes a first step upon the path along which Poincaré twenty-five years later made so great an advance, when he published his memoir, "Sur les propriétés arithmétiques des courbes algébriques" (Liouville, *Journal de Math.*, Series 5, Vol. VII (1901), p. 161). Poincaré considers Diophantine properties in connection with algebraic curves, while the geometrical meaning of equation (1) is a surface; thus Poincaré's work is not directly applicable here. But since the publication of his memoir it is impossible to overlook the importance of the geometrical meaning of results, and almost all that I have to say later depends upon geometrical ideas such as Poincaré here seized upon and grouped together systematically.

I cannot attempt to collect all the references to this problem since Euler's time, but I hope I have mentioned the most important. A paper upon Diophantine equations of degrees three and four in any number of variables which ought not to be passed over was published by Mr Robert Norrie in the Memorial Volume commemorating the five hundredth anniversary of St Andrew's University. The paper is of much interest, its scope is wide, and a large number of special problems are incidentally discussed. The solution of our problem which Mr Norrie gives is unfortunately not general; for it will be seen that his parameters $\lambda$ and $\mu$ cannot be rational unless one of the fractions such as $-(t+x)/(y+z)$ is the square of a rational quantity.

Some further results have been published even more recently by Prof. J. E. A. Steggall in the *Proc. Edinburgh Math. Soc.*, Vol. XXXIV (1916), p. 11. The paper contains interesting formulæ, notably a symmetrical solution of the equation

$$x^3 - u^3 = y^3 - v^3 = z^3 - w^3.$$

I wish also to express my thanks to Prof. Steggall for his kindness in allowing me to supplement my list of solutions from a table which he had calculated.

*Geometrical meaning of the Euler-Binet solution.* §§ 5—6.

5. The surface $t^3 + x^3 + y^3 + z^3 = 0$ is a special type of cubic surface, but it is not necessary to discuss its peculiarities at length. Of the 27 lines which lie on it three are real, viz.

$$t + x = y + z = 0; \quad t + y = z + x = 0; \quad t + z = x + y = 0; \quad \text{.................. (2)}$$

and lie in the plane

$$t + x + y + z = 0. \quad \text{........................................ (3)}$$

The other 24 lines are imaginary and have equations such as

$$t + \alpha x = 0, \quad z + \beta y = 0,$$

where $\alpha^3 = 1$ and $\beta^3 = 1$.

Consider two of these lines for which (i) $\alpha = \beta = \omega$, (ii) $\alpha = \beta = \omega^2$, where $\omega$ is one of the imaginary cube roots of unity. The equations

$$t + \omega x = h(z + \omega y), \quad t + \omega^2 x = k(z + \omega^2 y), \quad \text{.......................... (4)}$$

represent separately two planes passing each through one of the two lines: in combination they represent for different values of $h$ and $k$ any straight line which meets the two imaginary lines of the cubic surface. Such a line intersects the surface in a third point, whose coordinates must be expressible rationally in terms of $h$ and $k$. Conversely, through every point of the surface one such line can be drawn; thus the coordinates of every point of the surface can be thus expressed for proper values of $h$ and $k$. [An exception must be made for points which lie on either of the two chosen imaginary lines of the surface, or on any of the five imaginary lines of the surface which meet them both.] In fact by combining (4) with the equation of the surface we find

$$t + \omega x = h\,(z + \omega y),$$
$$t + \omega^2 x = k\,(z + \omega^2 y),$$
$$hk\,(t + x) + (y + z) = 0,$$

three linear equations which express $t$, $x$, $y$, $z$ as proportional to algebraic functions of $h$ and $k$. The functions contain the complex quantity $\omega$, so that real values of $t$, $x$, $y$, $z$ do not correspond to real values of the parameters $h$ and $k$. If we replace $h$ and $k$ by new parameters $f$ and $g$, $h$ and $k$ being equal to $f \pm ig$, real values of the parameters and of the variables will correspond. For application to the Diophantine equation a further condition is necessary, or at least extremely desirable, viz., that *rational* values of the variables and the parameters should correspond. This is effected by the introduction of parameters $p$ and $q$ in place of $h$ and $k$, such that

$$h = p \pm iq\sqrt{3}, \quad k = p \mp iq\sqrt{3};$$

and thus we are led to the Euler-Binet solution.

6. It is now clear that the problem of solving equation (1) in integers is almost identical with that of expressing the coordinates of points of the cubic surface (1) in terms of parameters, or in geometrical phraseology of establishing a one to one correspondence between the points of the surface and those of a plane. The Diophantine problem further demands that points of the surface and the plane whose coordinates are rational numbers should if possible correspond. This geometrical problem had been successfully attacked by Clebsch and by Cremona so far back as the year 1870; and in Geometry it has long been a familiar fact that the coordinates of the points of the general cubic surface are proportional to homogeneous polynomials of the *third* degree in three parameters. This result can be adapted to give a solution of the Diophantine equation of lower order in the parameters than that of Euler.

There are certain objections to the Euler-Binet formulae which are obvious, and others which are recognized as soon as the attempt is made to use the formulae for systematic calculation. It is not claimed that the new formulae are entirely free from the same faults, but they are affected to a much smaller extent. Thus it is undoubtedly more troublesome to work with two parameters which may assume any rational values than with three parameters which may be restricted to integer values (provided that the three parameters occur in a natural unforced manner). Again it is a fault in the Euler-Binet method that although the given equation is symmetrical in the four quantities $t$, $x$, $y$, $z$, the solution is completely unsymmetrical. This is really serious in practical calculation, for it means that any set of four numbers which satisfy the equation will reappear over and over again, each time in a different order, for other values of the parameters. The waste of time and labour so caused is very great: by the new formulae it can be avoided almost entirely. The Euler and the Euler-Binet formulae will furnish any number of illustrations of four numbers whose cubes have zero sum; but so far as I can discover they have never been used to calculate such sets of numbers systematically. In fact I cannot find that any table of solutions has been published.

*Parameter equations of the type due to Clebsch and Cremona.* §§ 7—11.

7. The coordinates of points of a cubic surface can be readily expressed in terms of three parameters if we can write the equation of the surface as a determinant of three rows, in which each element is a linear function of the coordinates. To do this with the given equation it is convenient to replace $t, x, y, z$ by new variables $T, X, Y, Z$. Let

$$2T = t + x + y + z, \quad 2Y = t - x + y - z,$$
$$2X = t + x - y - z, \quad 2Z = t - x - y + z,$$

so that

$$2t = T + X + Y + Z, \quad 2y = T - X + Y - Z,$$
$$2x = T + X - Y - Z, \quad 2z = T - X - Y + Z,$$

and

$$t - T = X - x = Y - y = Z - z = s,$$

where

$$2s = t - x - y - z = -T + X + Y + Z.$$

The last result gives the values of $t, x, y, z$ rapidly when $T, X, Y, Z$, are known and vice versa.

In terms of $T, X, Y, Z$, equation (1) becomes

$$T^3 + 3T(X^2 + Y^2 + Z^2) + 6XYZ = 0, \quad \text{.............................(2)}$$

and this may be written in the form

$$\begin{vmatrix} T, & 3Z, & -3Y \\ -Z, & T, & 3X \\ Y, & -X, & T \end{vmatrix} = 0. \quad \text{....................................(3)}$$

Now, if $T, X, Y, Z$ satisfy (3) it is both necessary and sufficient that quantities $a, b, c$ should exist such that

$$Ta + 3Zb - 3Yc = 0,$$
$$-Za + Tb + 3Xc = 0,$$
$$Ya - Xb + Tc = 0.$$

Solving these linear equations we can express both $T, X, Y, Z$ in terms of $a, b, c$, and $a, b, c$ in terms of $T, X, Y, Z$. Thus

$$\left.\begin{aligned} nT &= -6abc, \\ nX &= a(a^2 + 3b^2 + 3c^2), \\ nY &= b(a^2 + 3b^2 + 9c^2), \\ nZ &= c(3a^2 + 3b^2 + 9c^2), \end{aligned}\right\} \quad \text{....................................(E)}$$

and

$$\left.\begin{aligned} a : b : c &:: T^2 + 3X^2 : 3XY + TZ : ZX - TY, \\ &:: 3(XY - TZ) : T^2 + 3Y^2 : 3YZ + TZ, \\ &:: 3(3ZX + TY) : 3(YZ - TX) : T^2 + 3Z^2. \end{aligned}\right\} \quad \text{............................(F)}$$

[We could pass from these rational but unsymmetrical formulae to others which are perfectly symmetrical in $X, Y, Z$, but contain surds, by replacing $a, b, c$, by $g^2a, gb, c$, where $g^3 = 3$.]

Thus to every set of values of $T, X, Y, Z$ satisfying (2) corresponds a set of values of $a : b : c$, and vice versa*. To rational values correspond rational values. We confine our attention as has been stated above, to sets of integer values of $t, x, y, z$, which have no factor common to all.

* There are certain exceptions to this statement, as is known in geometry; but they do not concern us, the values being imaginary. Cf. § 5.

It follows that we may obtain every such solution of the Diophantine equation by giving integer values (having no common factor) to $a$, $b$, $c$ in equations (E), calculating $t$, $x$, $y$, $z$, and removing common factors.

8. We will now suppose that a set of values of $t$, $x$, $y$, $z$, which satisfy (1) has been found. Other sets may be derived by permuting these values in any way, and by changing the signs of all. As was remarked in § 6, it is wasted labour if we obtain these independently from another set of values of the parameters $a$, $b$, $c$. We must consider how to avoid this as far as possible.

Since it is permissible to change the signs of $t$, $x$, $y$, $z$ let it be agreed that the *numerically greatest* member of any set of values shall have a *positive* sign; the other members will then be either all negative or two negative and one (the smallest in absolute value) positive. Further let this numerically greatest member of any set be denoted by $t$. It is easy to see that $X$, $Y$, $Z$ are all positive, and therefore by (2) $T$ is negative. The values of $a$, $b$, $c$ must clearly all be positive. Hence

I. *It is unnecessary to consider any but positive values of the parameters a, b, c.* Negative values of $a$ or $b$ or $c$ only lead to the same values of $t$, $x$, $y$, $z$ in a different order.

In the known set of values, the symbol $t$ has been assigned to a special member, but the symbols $x$, $y$, $z$ may be allotted to the remaining members in six ways. Any interchange of $x$, $y$, $z$ causes the same interchange of $X$, $Y$, $Z$. Now if the given values $t$, $x$, $y$, $z$ are obtained by equations (E) from the positive values $a$, $b$, $c$ of the parameters, and if $t$, $y$, $z$, $x$ are derived from values $a'$ $b'$, $c'$, and $t$, $z$, $x$, $y$ from $a''$, $b''$, $c''$, equations (F) shew that

$$a' : b' : c' :: T^2 + 3Y^2 : 3YZ + TX : XY - TZ$$
$$:: b : c : \tfrac{1}{3}a :: 3b : 3c : a$$

and

$$a'' : b'' : c'' :: T^2 + 3Z^2 : 3ZX + TY : YZ - TX$$
$$:: c : \tfrac{1}{3}a : \tfrac{1}{3}b :: 3c : a : b.$$

It is unnecessary to consider more than one of the three sets of parameters $a$, $b$, $c$; $a'$, $b'$, $c'$; $a''$, $b''$, $c''$: and if we stipulate that

II. *The first parameter a must not be a multiple of* 3, we rule out two of the three, retaining the simplest. For clearly if $a$ be divisible by three we may reject the system of parameters $a$, $b$, $c$ in favour of the simpler system $b$, $c$, $\tfrac{1}{3}a$, which will lead to the same values of $t$, $x$, $y$, $z$ in a different order.

9. Finally (still supposing that the values $t$, $x$, $y$, $z$ are obtained from the parameters $a$, $b$, $c$), we have to consider the values of the parameters which give the solutions $(t, x, z, y)$ $(t, z, y, x)$ $(t, y, x, z)$. In these three the last three letters are interchanged cyclically as in those discussed in section 8. Hence one of the three is derived from positive values $(a_1, b_1, c_1)$ of the parameters, in which $a_1$ is not a multiple of 3.

But the relations between the parameters $(a, b, c)$ and $(a_1, b_1, c_1)$ are by no means simple. When either set is given the other is determinate and is given by equations which can be written down explicitly. For by interchange of $X$ and $Y$ in the first of equations (F) we find

$$a_1 : b_1 : c_1 :: T^2 + 3Y^2 : 3XY + TZ : YZ - TX$$
$$:: b\,[(a^2 + 3b^2 + 9c^2)^2 + 12a^2c^2] : a\,[(a^2 + 3b^2 + 3c^2)^2 + 12b^2c^2] : c\,[3\,(a^2 + b^2 + 3c^2)^2 + 4a^2b^2]$$

after a little simplification. These relations are too complicated to be helpful. The simplest procedure seems to be to ignore them and state the theorem we have established in the form

*Every set of four integers (having no factor common to all) which satisfy the Diophantine equation*

$$t^3 + x^3 + y^3 + z^3 = 0, \qquad (1)$$

*may be obtained from equations* (E) *in two ways from sets of positive values of the parameters* $a, b, c$, *the parameter* $a$ *not being a multiple of* 3.

**10.** The values of the parameters which correspond to a given solution are obtained from equations (F), or from those equations with two of the letters $X, Y, Z$ interchanged. Should the value of $a$ turn out to be a multiple of 3, we replace the parameters $a, b, c$ by $b, c, \frac{1}{3}a$ as explained above: this, however, will never be necessary if a little care is taken. For equations (E) shew that $T$ and $Z$ are always and that $X$ is never divisible by 3, while $Y$ is or is not divisible by 3 according as $b$ is or is not divisible by 3. Thus when $T, X, Y, Z$ are known, if only one of $X, Y, Z$ is a multiple of 3, we call that one $Z$; if only one is not a multiple of 3, we call that one $X$. In each case the other symbols may be assigned in two ways which lead directly to two sets of values of the parameters, in which $a$ is not a multiple of 3.

[It will be observed that, if $b$ is a multiple of 3, none of the numbers $t, x, y, z$ is a multiple of 3; but when $b$ is not a multiple of 3, two of them (either $t$ and $z$ or $x$ and $y$) are multiples of 3. Similarly, if none or two of the parameters are even, all the numbers $t, x, y, z$ are odd.]

To take an example, let us assume that

$$a = 2, \quad b = 3, \quad c = 1.$$

Then $$T = -36, \quad X = 68, \quad Y = 120, \quad Z = 48; \quad s = 136;$$

and, solving for $t, x, y, z$, we have, after rejecting the common factor 4,

$$t = 25, \quad x = -17, \quad y = -4, \quad z = -22,$$

shewing that $$25^3 = 22^3 + 17^3 + 4^3.$$

Again, let us find the two sets of values of the parameters which lead to the solution

$$41^3 = 40^3 + 17^3 + 2^3.$$

Here $$t = 41, \quad x, y, z, = -40, -17, -2,$$

$$T = -9, \quad X, Y, Z, = 10, \quad 33, \quad 48.$$

We must take either

$$\text{(i)} \quad T = -9, \quad X = 10, \quad Y = 33, \quad Z = 48,$$

or $$\text{(ii)} \quad T = -9, \quad X = 10, \quad Y = 48, \quad Z = 33.$$

Taking the first of equations (F), but making an obvious simplification by dividing by 3, we have

$$a : b : c :: X^2 + \tfrac{1}{3}T^2 : XY + \tfrac{1}{3}TZ : \tfrac{1}{3}ZX - \tfrac{1}{3}TY,$$

$$\text{(i)} \quad :: 127 : 186 : 259,$$

$$\text{(ii)} \quad :: 127 : 381 : 254 :: 1 : 3 : 2.$$

Thus we have two sets of values of $a, b, c$; either

$$(1, 3, 2); \quad \text{or} \quad (127, 186, 259).$$

**11.** If the formulae (E) are to be used to calculate systematically the solutions of our equation, it is not a little disconcerting to find that a simple solution should be derived from such large values of the parameters. Yet the form of the equations of §9 connecting $a_1, b_1, c_1$, and $a, b, c$ shews that it must constantly happen that the first set of parameters may be small and the second set quite large numbers. Large parameters can never lead to simple values of $T, X, Y, Z$ and $t, x, y, z$ except when the expressions in equations (E) contain large common factors which can be cancelled out. It will be seen that such common factors must be factors either of $a$ and $b^2+3c^2$, or of $b$ and $a^2+3c^2$ or of $c$ and $a^2+3b^2$. When for example $a=127$, $b=186$, $c=259$, it will be found that 127 is a factor of $b^2+3c^2$, 259 is a factor of $a^2+3b^2$, and 62 is a common factor of $b$ and $a^2+3c^2$: hence there is a vast reduction in the values of $T, X, Y, Z$ and $t, x, y, z$ given by equations (E). [It is an essential feature of the parametric equations of Clebsch and Cremona that the cubic functions of $a, b, c$ should all vanish for six sets of values of $a, b, c$. In our equations these values are given by $a=0$, $b^2+3c^2=0$; $b=0$, $a^2+3c^2=0$; $c=0$, $a^2+3b^2=0$; hence the above rule regarding common factors. A table of all numbers less than 1000 which are of the form $a^2+3b^2$ is included in the paper of Euler referred to in § 1.]

Although every solution of the equation can be obtained from equations (E), and although the repetitions and other difficulties of the older method are to a great extent avoided, yet it is never possible without a careful scrutiny to feel confident that all solutions of a certain type (e.g. all solutions in which no number exceeds a certain limit, 50 or 100, or all solutions in which the two largest numbers differ by unity) have been found. For this reason it is convenient to have in addition certain simple methods (again suggested by the geometry of the cubic surface) by which new solutions may be deduced from those that have been obtained.

*Methods of deriving new solutions from a known solution.* §§ 12—13.

12. Every solution corresponds to a point on the cubic surface whose coordinates are integers, and by permuting the coordinates we have a family of 24 points. Now, if two points are taken on a surface of order three, the line joining them cuts the surface in a third point, whose coordinates are determinate, and are rational when those of the two known points are rational. For example the points $(6, -5, -4, -3)$ and $(6, -4, -3, -5)$ lie on the surface, and the coordinates of any point of the line joining them may be written

$$6\lambda+6\mu,\quad -5\lambda-4\mu,\quad -4\lambda-3\mu,\quad -3\lambda-5\mu.$$

If this point lie on the surface, the sum of the cubes of these four expressions will vanish, and therefore

$$\lambda\mu\,(23\lambda+25\mu)=0.$$

Giving $\lambda$ and $\mu$ the values 25 and $-23$ respectively, we find a new solution $(12, -33, -31, 40)$.

Similarly from the two solutions $(9, -8, -6, -1)$ $(12, -9, -10, 1)$ we derive a new solution $(-21, -43, 88, -84)$. From the two $(9, -8, -6, -1)$ $(12, -10, -9, 1)$ we derive only the trivial solution $(3, -2, -3, 2)$.

Other methods of deducing solutions will be found in the books or papers to which reference was made in §§ 1—4. Thus Vieta (see Heath's *Diophantus*, p. 102 footnote), starting from a known solution $x^3+y^3+a^3+b^3=0$, regards $a$ and $b$ as constants, $x$ and $y$ as variables: from this point of view the equation represents a plane cubic curve on which one point $(x, y)$ is known. Vieta draws the tangent line at that point and so finds another point (where the tangent cuts the curve) whose coordinates give a new solution, rational but usually not integral, of the equation, i.e. a solution in which two of the unknowns $a, b$ have the same values as in the known solution. Again Norrie has a general method, applicable to cubic equations in any number of variables, which for our equation amounts to this. Knowing a solution of equation (1) he knows a point on the cubic surface. He constructs the tangent plane at that point, and joins the known point to an arbitrary point in the tangent plane. The joining line cuts the surface in another determinate point. Thus from one known solution ($t_0$, $x_0$, $y_0$, $z_0$) he deduces an infinity of solutions ($t$, $x$, $y$ $z$) all satisfying both the given equation (1) and

$$t_0^2 t + x_0^2 x + y_0^2 y + z_0^2 z = 0.$$

The application of this method turns out to be somewhat troublesome.

13. *From any known solution it is possible to derive all solutions in which any one of the fractions*

$$(t+x)/(y+z), \quad (t+y)/(z+x), \quad (t+z)/(x+y),$$

*has the same value as in the known solution.* This appears to be the simplest method both in theory and practice of deriving new solutions. If we know a point ($t$, $x$, $y$, $z$) on the surface and join it to an arbitrary point ($-h$, $h$, $-k$, $k$) on the line

$$t + x = 0, \quad y + z = 0,$$

the point at which the joining line cuts the surface has coordinates

$$\left.\begin{aligned} &t - \theta h, \ x + \theta h, \ y - \theta k, \ z + \theta k, \\ \text{where} \qquad &\theta = \frac{(t^2 - x^2) h + (y^2 - z^2) k}{(t+x) h^2 + (y+z) k^2}. \end{aligned}\right\} \quad \text{.................................(G)}$$

By taking any rational value of $h/k$, i.e. by giving to $h$ and $k$ any integer values (positive or negative) prime to one another, we derive a new solution. Should $\theta$ be an integer, the values of $t+x$ and $y+z$ in the new solution are either equal to or less than their values in the known solution. By interchange of $x$, $y$ and $z$ a triple infinity of solutions is derived from any known solution.

Formulae such as this which give an infinity of solutions of the equation define curves on the cubic surface, and when the coordinates are expressed as rational algebraic functions of a parameter ($h/k$ in equations G) the curves must be unicursal. The points of a plane section of the surface, a cubic curve without a node, cannot in general be so expressed; but if the plane touch the surface they can be so represented, as Norrie shews. The simplest curves on the surface are conics cut out by planes which pass through one of the lines of the surface, and these are what equations (G) define. Since they are conics, the coordinates are proportional to *quadratic* functions of $h$, $k$; all other curves require functions of a higher order than the second.

[Another family of curves giving systems of solutions may be briefly noticed. If in equations (E) we put $b = c$, or $a = b + c$, or if we impose any linear relation $a = Bb + Cc$ upon the parameters, we obtain a curve on the surface and a system of solutions. The curves will be found to be twisted curves of order three.]

### *Equations* (E) *and* (G) *considered as standard solutions.* § 14.

14. The view which the writer of this paper wishes to put forward is that equations (E), combined with the supplementary system (G), should be regarded as the standard solution of the Diophantine equation. The former are based upon the accepted parametric equations of a cubic surface; they include all solutions: and (under the two simple rules of § 8) avoid useless repetitions almost entirely. They are however open to the objection (though in a less degree than the old

solutions) that if they are used in calculating solutions there is always a possibility that some simple solution may be overlooked, because both the sets of parameters that lead to that solution may be large. [It is not suggested that to tabulate solutions is the final object of the investigation.] For this reason it is advisable to retain equations (G) as a supplementary set of formulae. They represent the simplest curves (conics) which lie on the cubic surface and wholly cover it, and are the only such curves expressible by homogeneous quadratic functions of two parameters. The formulae are very easily applied as soon as a solution of the Diophantine equation has been obtained from equations (E) or by any other method. Moreover it will be found that when the fractions $(t+x)/(y+z)$, etc., are reduced to their lowest terms, both numerator and denominator are numbers of the particular form $m^2+3n^2$ (which Euler has tabulated, up to 1000, in the paper referred to in § 1). The effect of this is that the number of cases of equations (G) that arise is far less than would at first sight be expected. Table I of § 19 shews this very plainly, by the frequent reappearances of the fractions in Column III.

*Further consideration of equations* (G). §§ 15—16.

15. If we again adopt the convention of § 8 that $t$ shall be the numerically greatest of the four numbers $t, x, y, z$ and shall be positive, the three fractions

$$(t+x)/(y+z), \quad (t+y)/(z+x), \quad (t+z)/(x+y)$$

have positive numerators and negative denominators; also each is numerically less than unity. For, expressed in terms of $a, b, c$ (which are positive),

$$\frac{t+x}{y+z} = \frac{T+X}{T-X} = -\frac{a^2+3(b-c)^2}{a^2+3(b+c)^2},$$

$$\frac{t+y}{z+x} = \frac{T+Y}{T-Y} = -\frac{(a-3c)^2+3b^2}{(a+3c)^2+3b^2},$$

$$\frac{t+z}{x+y} = \frac{T+Z}{T-Z} = -\frac{(a-b)^2+3c^2}{(a+b)^2+3c^2}.$$

Further, the numerators and denominators of the fractions on the right are all of them numbers of the special form $m^2+3n^2$, and, by the well known theorem concerning the factors of such numbers, the numerators and denominators are still of this form when the fractions are reduced to their lowest terms. This result is used by Euler, and may be proved also from the fact that, if

$$t^3+x^3+y^3+z^3=0, \quad \text{......................................(1)}$$

$$\frac{t+x}{y+z} = -\frac{y^2-yz+z^2}{t^2-tx+x^2} = -\frac{(y+z)^2+3(y-z)^2}{(t+x)^2+3(t-x)^2}.$$

In Euler's solution (A), quoted in § 1, we see that

$$\frac{t+x}{y+z} = -\frac{s^2+3r^2}{p^2+3q^2}.$$

In order to obtain simple solutions, Euler selects from his table two values for $s^2+3r^2$ and $p^2+3q^2$ which have a fairly large common factor, and which are obtained from several sets of values of $s, r, p, q$; he then follows out the various cases and so obtains several solutions for all of which $(t+x)/(y+z)$ has the same value. In his first example he selects values 19 and 76 and so gets seven solutions for which $(t+x)/(y+z)$ is $-1/4$; in his second example he selects values 28 and 84, and so gets eight solutions in which $(t+x)/(y+z)$ is $-1/3$. The solutions he obtains are not the simplest for which $(t+x)/(y+z)$ has these values.

16. To illustrate the use of equations (G) we will take the simplest solution of the Diophantine equation and derive the three systems of solutions.

$$t=6, \quad x=-5, \quad y=-4, \quad z=-3.$$

Applying the method of § 13, we have the following results. The general solution for which

(i) $(t+x)/(y+z)=-1/7$ is

$$t=6-\theta h, \quad x=-5+\theta h, \quad y=-4-\theta k, \quad z=-3+\theta k,$$

where $$\theta=(11h+7k)/(h^2-7k^2);$$

(ii) $(t+y)/(z+x)=-1/4$ is

$$t=6-\theta h, \quad y=-4+\theta h, \quad z=-3+\theta k, \quad x=-5-\theta k,$$

where $$\theta=(10h+8k)/(h^2-4k^2);$$

(iii) $(t+z)/(x+y)=-1/3$ is

$$t=6-\theta h, \quad z=-3+\theta h, \quad x=-5-\theta k, \quad y=-4+\theta k,$$

where $$\theta=(9h+3k)/(h^2-3k^2).$$

It will be observed that until $h$ and $k$ receive definite values, it is not possible to say which of the four numbers is numerically greatest. The rule of § 8 with regard to $t$ must be abandoned for the moment.

From (i) we learn that there is an infinite sequence of solutions in which $t+x=1$, $y+z=-7$, derived from integer values of $\theta$. For an infinity of values of $h$ and $k$ (positive and negative) make $h^2-7k^2=1$, and each of these gives such a solution. [Other values of $h$ and $k$ make $h^2-7k^2=2$, or $-3$, and these also give infinite sequences of solutions in which $t+x=1$, $y+z=-7$.] So from (iii) we can derive an infinite sequence of solutions in which $t+x=3$, $y+z=-9$.

But in (ii) the case is different, since the coefficient of $k^2$ is a perfect square. Here

$$\theta=7/(h-2k)+3/(h+2k),$$

and there are only a finite number of integer values of $h$ and $k$ that make $\theta$ integral. The number of solutions for which $(t+y)/(z+x)=-1/4$ is of course infinite, and all can be derived from (ii), by giving $h$ and $k$ integer values (positive or negative) prime to one another.

In the application of equations (G) the same difficulty arises which was discussed in § 8, that when a solution $(t'x'y'z')$ has been derived from the solution $(t, x, y, z)$ from certain values of $h, k$, certain other values must lead to the solutions $(t'x'z'y')$ $(x't'y'z')$ $(x't'z'y')$, which clearly have the same value of $(t+x)/(y+z)$. As in § 8 there is a simple rule by which the number of repetitions can be reduced from 3 to 1, so that the waste of labour that might be due to this cause is to a large extent avoided, viz., that if

$$(t+x)\,hh'+(y+z)\,kk'=0$$

the parameters $(h', k')$ lead to the same set of values as $(h, k)$ in a different order.

*Examples of the formulae and methods.* §§ 17—19.

17. In conclusion I will give a few applications of these formulae and methods. It is always to be understood that in any known numerical solution $(t, x, y, z)$, $t$ is positive and is numerically greater than $x$ or $y$ or $z$. In an algebraic formula it may not be possible to say which number is greatest. No assumption is made as to the relative magnitudes of $x, y, z$. When the parameters $a, b, c$ are employed, it is to be understood that they are positive and that $a$ is not a multiple of 3.

*Example I. To find the general solution of the equation for which* $(t+x)/(y+z) = -1/4$.

In order to use the formulae (G) we must know a special solution of the kind. This has already been found, viz. $(6, -4, -5, -3)$, and the formulae were given in § 16 (ii). It was stated there that since $-(t+x)/(y+z)$ has a rational square root the characteristics of the system of solutions differ in some respects from what is usual. The formulae of § 16 for the four numbers $t, x, y, z$, viz.

$$6 - \theta h, \quad -4 + \theta h, \quad -5 - \theta k, \quad -3 + \theta k,$$
$$\theta = (10h + 8k)/(h^2 - 4k^2),$$

can be simplified by the substitution

$$h + 2k : h - 2k :: f : g.$$

They become

$$4fg \pm 2(7f^2 + 3g^2), \quad -16fg \pm (7f^2 - 3g^2).$$

We note that when $f$ and $g$ are both odd a factor 4 cancels out, and that when $f$ is a multiple of 3, a factor 3 cancels out. Also since parameters $(f', g')$ lead to the same numbers as $(f, g)$ when

$$\text{(i)}\ f'/g' = -f/g, \quad \text{(ii)}\ 7ff' \pm 3gg' = 0,$$

it follows that we can avoid repetitions entirely if we take only positive values of $f$ and $g$, the latter not being divisible by 7. It is now not difficult to obtain the following twelve solutions, which I believe to be the complete system when no number exceeds 100.

Table of solutions when $(t+x)/(y+z) = -1/4$.

| $f$ | $g$ | $t$ and $x$ | | $y$ and $z$ | |
|---|---|---|---|---|---|
| 1 | 1 | 6, | −4 | −5, | −3 |
| 1 | 2 | 46, | −30 | −37, | −27 |
| 1 | 3 | 20, | −14 | −17, | −7 |
| 1 | 5 | 46, | −36 | −37, | −3 |
| 2 | 1 | 70, | −54 | −57, | −7 |
| 3 | 1 | 12, | −10 | −9, | +1 |
| 3 | 2 | 58, | −42 | −49, | −15 |
| 3 | 4 | 90, | −58 | −69, | −59 |
| 3 | 5 | 28, | −18 | −21, | −19 |
| 3 | 11 | 82, | −60 | −69, | −19 |
| 5 | 1 | 94, | −84 | −63, | +23 |
| 9 | 1 | 98, | −92 | −59, | +35 |

18. *Calculation of solutions by equations* (E). It is easy to discover any number of solutions of the Diophantine equation by means of equations (E), but it becomes clear at the outset that the lowest values of the parameters do not lead to the simplest solutions. Thus the values $a = b = c = 1$ give the solution $(29, -27, -15, -11)$, while the simplest solutions of all $(6, -5, -4, -3)$ and $(9, -8, -6, -1)$ are derived from the values $(2, 1, 1)$ and $(1, 2, 1)$ of $(a, b, c)$. The explanation is easily seen, and when the parameters are known it is possible to foretell, without actually going through the calculations, whether the solution will be a simple one or not. The values of $(T, X, Y, Z)$ of the formulae in (E) are simplified when they contain a common factor, which can be cancelled. The factor cannot be 3, since $a$ and therefore $X$ is not a multiple of 3. For all odd factors, the rule given in §11 is valid, viz. factors which can be cancelled out must be common to one of the three following pairs of numbers

$$\text{(i) } a \text{ and } b^2 + 3c^2, \quad \text{(ii) } b \text{ and } a^2 + 3c^2, \quad \text{(iii) } c \text{ and } a^2 + 3b^2.$$

Such factors must be of the form $m^2 + 3n^2$.

It remains to consider when it is possible to reject a factor 2 or a power of 2. This case cannot be included in the general statement on account of the irregular rôle which 2 plays in connection with the factors of numbers $m^2 + 3n^2$; and it is confused by the multiplier 2 which occurs in the equations connecting $(T, X, Y, Z)$ and $(t, x, y, z)$. When two parameters are odd and one is even, a factor 4 can be rejected from $(T, X, Y, Z)$ leaving quotients of which two are even and two odd. The values of $t, x, y, z$ are then found by the equations; two of them are even and two odd. When two parameters are even and one odd, or when all three are odd, $T$ is even, $X, Y, Z$ are odd. The values of $t, x, y, z$ as given by the equations are not integers, and it is necessary to *introduce a factor* 2. When this has been done, integer values, all of them *odd*, are found for $t, x, y, z$.

Thus a set of parameters $a, b, c$, whose sum is even, gives values of $t, x, y, z$ which are, roughly speaking, only one-eighth of those derived from parameters of nearly the same values whose sum is odd. It will be found that quite low sets of parameters, e.g. $(2, 1, 2)$, $(2, 2, 1)$, $(1, 3, 1)$, $(1, 1, 3)$, $(1, 3, 3)$ lead to solutions containing numbers which exceed 100. On the other hand the set $(7, 2, 1)$, where $a$ and $c$ are odd, $b$ is even, and in addition $a = b^2 + 3c^2$, leads to the very simple solution $(12, -10, -9, 1)$; a factor 28 is common to $T, X, Y, Z$. The same factor 28 could be rejected for the sets $(2, 7, 1)$, $(2, 1, 7)$, $(7, 1, 4)$, $(7, 3, 2)$.

It is thus possible to foretell what factors will cancel out, as soon as a set of parameters has been chosen. The odd factors are all those contained in the three greatest common measures of $a$ and $b^2 + 3c^2$, of $b$ and $a^2 + 3c^2$, and of $c$ and $a^2 + 3b^2$; the power of 2 is given by the rule above, and depends only on whether the sum of the parameters is odd or even.

19. *Tables of integers which satisfy the equation*

$$t^3 + x^3 + y^3 + z^3 = 0.$$

Below are given 19 sets of integers $t, x, y, z$, none greater than 50, which satisfy the equation, which I believe to be the total number of such solutions. In a second column are shewn the values of the three fractions such as $-(t + x)/(y + z)$ corresponding to each solution, and in a third column the values of the parameters $a, b, c$ which lead to the solution. In Table II some further solutions in numbers less than 100 are shewn; the list is possibly complete; there may be omissions, but not many. In III a statement made in §15 is illustrated, and a note on a method of numerical calculation is added in IV.

I. Table of sets of numbers less than 50 which satisfy

$$t^3 + x^3 + y^3 + z^3 = 0.$$

| I. $t$ | II. $x, y, z$ | | | III. $-(t+x)/(y+z)$ etc. | | | IV. $a$ | $b$ | $c$ |
|---|---|---|---|---|---|---|---|---|---|
| 6 | − 5, | − 4, | − 3 | 1/7, | 1/4, | 1/3 | 2, | 1, | 1 |
| 9 | − 8, | − 6, | − 1 | 1/7, | 1/3, | 4/7 | 1, | 2, | 1 |
| 12 | − 10, | − 9, | + 1 | 1/4, | 1/3, | 13/19 | 7, | 2, | 1 |
| 16 | − 15, | − 9, | + 2 | 1/7, | 7/13, | 3/4 | 1, | 1, | 2 |
| 19 | − 18, | − 10, | − 3 | 1/13, | 3/7, | 4/7 | 4, | 1, | 1 |
| 20 | − 17, | − 14, | − 7 | 1/7, | 1/4, | 13/31 | 7, | 3, | 2 |
| 25 | − 22, | − 17, | − 4 | 1/7, | 4/13, | 7/13 | 2, | 3, | 1 |
| 27 | − 24, | − 19, | + 10 | 1/3, | 4/7, | 37/43 | 2, | 7, | 1 |
| 28 | − 21, | − 19, | − 18 | 7/37, | 3/13, | 1/4 | 13, | 10, | 7 |
| 29 | − 27, | − 15, | − 11 | 1/13, | 7/19, | 3/7 | 1, | 1, | 1 |
| 34 | − 33, | − 15, | + 2 | 1/13, | 19/31, | 3/4 | 7, | 1, | 4 |
| 34 | − 33, | − 16, | + 9 | 1/7, | 3/4, | 43/49 | 13, | 1, | 2 |
| 39 | − 36, | − 26, | + 17 | 1/3, | 13/19, | 28/31 | 26, | 7, | 1 |
| 40 | − 33, | − 31, | + 12 | 7/19, | 3/7, | 13/16 | 1, | 4, | 1 |
| 41 | − 33, | − 32, | − 6 | 4/19, | 3/13, | 7/13 | 5, | 2, | 1 |
| 41 | − 40, | − 17, | − 2 | 1/19, | 4/7, | 13/19 | 1, | 3, | 2 |
| 44 | − 41, | − 23, | − 16 | 1/13, | 7/19, | 7/16 | 4, | 3, | 1 |
| 46 | − 37, | − 30, | − 27 | 3/19, | 1/4, | 19/67 | 16, | 13, | 7 |
| 46 | − 37, | − 36, | − 3 | 3/13, | 1/4, | 43/73 | 10, | 19, | 7 |

II. Other solutions in integers which do not exceed 100.

| | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 51, | 43, | 38, | + 12 | 75, | − 66, | 43, | − 38 | 89, | − 86, | − 40, | − 17 |
| 53, | − 44, | − 34, | − 29 | 76, | − 69, | − 48, | + 5 | 89, | − 86, | − 41, | + 2 |
| 53, | − 50, | − 29, | + 8 | 76, | − 72, | − 33, | − 31 | 90, | − 69, | − 59, | − 58 |
| 54, | − 53, | − 19, | − 12 | 76, | − 73, | − 38, | + 17 | 90, | − 87, | − 38, | − 25 |
| 55, | − 54, | − 24, | + 17 | 80, | − 71, | − 54, | + 15 | 93, | − 92, | − 30, | + 11 |
| 58, | − 49, | − 42, | − 15 | 81, | − 74, | − 48, | − 25 | 93, | − 85, | − 54, | − 32 |
| 58, | − 57, | − 22, | + 9 | 82, | − 75, | − 64, | + 51 | 94, | − 84, | − 63, | + 23 |
| 60, | − 59, | − 22, | + 3 | 82, | − 69, | − 60, | − 19 | 96, | − 90, | − 53, | − 19 |
| 67, | − 54, | − 51, | − 22 | 84, | − 75, | − 53, | − 28 | 96, | − 93, | − 59, | + 50 |
| 67, | − 58, | − 51, | + 30 | 85, | − 64, | − 61, | − 50 | 97, | − 79, | − 69, | − 45 |
| 69, | − 61, | − 36, | − 38 | 87, | − 78, | − 55, | − 26 | 97, | − 96, | − 33, | + 20 |
| 69, | − 61, | − 56, | + 42 | 87, | − 79, | − 48, | − 38 | 97, | − 90, | − 66, | − 47 |
| 70, | − 57, | − 54, | − 7 | 87, | − 79, | − 54, | − 20 | 98, | − 92, | − 59, | − 35 |
| 71, | − 70, | − 23, | − 14 | 88, | − 84, | − 43, | − 21 | 98, | − 89, | − 63, | − 24 |
| 72, | − 65, | − 39, | − 34 | 88, | − 86, | − 31, | − 25 | | | | |

III. The infinite sequences of solutions referred to in § 16, in which

$$t + x = 1, \quad y + z = -7,$$
$$\left.\begin{array}{l} t \text{ and } x = \quad 6 - \lambda h, \ -5 + \lambda h, \\ y \text{ and } z = -4 - \lambda k, \ -3 + \lambda k, \end{array}\right\}$$
$$\lambda = (11h + 7k)/(h^2 - 7k^2),$$

are as follows.

A. $h^2 - 7k^2 = 1$.

| | | | | |
|---|---|---|---|---|
| $h =$ | 8, | 127, | 2024, | ... |
| $k =$ | $\pm 3$, | $\pm$ 48, | $\pm$ 765, | ... |

B. $h^2 - 7k^2 = 2$.

| | | | | |
|---|---|---|---|---|
| $h =$ | 3, | 45, | 717, | ... |
| $k =$ | $\pm 1$, | $\pm$ 17, | $\pm$ 271, | ... |

C. $h^2 - 7k^2 = -3$.

| | | | | | |
|---|---|---|---|---|---|
| | $h =$ | 5, | 82, | 1307, | ... |
| | $k =$ | 2, | 31, | 494, | ... |
| and | $h =$ | 2, | 37, | 590, | ... |
| | $k =$ | $-1$, | $-$ 14, | $-$ 229, | ... |

In each sequence the values of $h$ and $k$ form a recurring series, the successive terms obeying the law

$$u_{n+1} = 16u_n - u_{n-1}.$$

IV. Another version of equations (G), adopted from Euler's memoir, is specially suitable for the numerical calculation of solutions. I have used it to check and complete Tables I and II.

If $(t + x)/(y + z) = a/(-b)$, $a$ and $b$ being prime to one another and $a < b$, and we write

$$t = ka + u, \quad x = ka - u, \quad y = -kb + v, \quad z = -kb - v,$$

then

$$3au^2 - 3bv^2 = (b^3 - a^3) k^2.$$

It is not necessary here to enter into the theory of this quadratic equation. If we find positive integer values of $u$, $v$, $k$ having no common factor, the values of $t$, $x$, $y$, $z$, have no common factor, except a 2 when $a$, $b$, $k$, $u$, $v$ are all odd. When values of $a$, $b$ and $k$ have been chosen, it is easy, with the help of a table of squares and a calculating machine, to find the values of $u$ and $v$ that give all the solutions in $t$, $x$, $y$, $z$ up to any imposed limit. I have done this for all values of $a$ and $b$ up to and including $b = 16$.

CAMBRIDGE: PRINTED BY J. B. PEACE, M.A., AT THE UNIVERSITY PRESS

# XX. *On Cyclical Octosection.*

By W. Burnside, M.A., F.R.S., Hon. Fellow of Pembroke College.

[*Received* 10 April, *Read* 3 May, 1920.]

The complete solution of the problem of cyclical quartisection was first given by V. A. Le Besgue in the *Comptes Rendus* [Vol. LI, p. 10 (1860)]. The result is given in the form that, if $\lambda$ is the sum of $\frac{1}{4}(p-1)$ distinct primitive $p$th roots of unity which sum takes just four distinct values, then $1+4\lambda$ is a root of the equation

$$[y^2 + \{1 - 2(-1)^{\frac{p-1}{4}}\} p]^2 - 4p[y - L]^2 = 0,$$

where $$p = L^2 + 4M^2, \quad L \equiv 1 \pmod{4}.$$

No proof is given. The only proof that I know of is one by P. Bachmann in the sixteenth chapter of his work on *Kreistheilung*. This proof appears to me to be quite unnecessarily complicated; and I have therefore established the formulae, so far as they are necessary for the problem of octosection, independently.

1. It may be convenient to recall those properties of an algebraic number-field of which use will be made. By a rational number is meant a fraction $\frac{p}{q}$, of which the numerator $p$ and denominator $q$ are ordinary integers (in the sense of elementary arithmetic) affected either with the positive or the negative sign. In the particular case in which $q$ is unity, the rational number is said to be a rational integer.

If $x$ satisfies an equation of the $n$th degree, the coefficients of which are rational numbers, and if it is impossible to express the left-hand side of the equation as the product of two factors in which the coefficients are rational numbers, the equation is said to be irreducible. The totality of the rational functions of $x$, which satisfies such an irreducible equation with rational coefficients, is called an algebraic number-field of order $n$.

If $y$ satisfies an equation of finite degree with rational integral coefficients and if the leading coefficient is $\pm 1$, $y$ is called an algebraic integer.

The principal property of an algebraic number-field that will here be made use of is the following:

In an algebraic number-field of order $n$, a set of $n$ algebraic integers $x_i$ $(i = 1, 2, \ldots, n)$ can be chosen (in an infinite number of ways), such that every algebraic integer belonging to the field can be expressed in the form

$$a_1 x_1 + a_2 x_2 + \ldots + a_n x_n,$$

where $a_1, a_2, \ldots, a_n$ are rational integers. The set of algebraic integers $x_i$ $(i = 1, 2, \ldots, n)$ is called an integral basis of the field.

*Notation.* The prime $p$ is congruent to unity (mod. 8); $g$ is a primitive root of the congruence $g^{p-1} \equiv 1$ (mod. $p$);

$\omega$ is a primitive $p$th root of unity;

$$\mu_i = \sum_{n=0}^{n=\frac{1}{8}(p-1)-1} \omega^{g^{8n+i-1}} \quad (i = 1, 2, \ldots, 8);$$

$\mu_i + \mu_{i+4} = \lambda_i$;

$\sqrt{\alpha}$, where $\alpha$ is a real positive quantity, is the positive square-root of $\alpha$.

2. The algebraic number-field defined by $\mu_1$ is of order 8. An integral basis for the field* is given by $\mu_1, \mu_2, \ldots, \mu_8$. The field defined by $\lambda_1$ is of order 4. An integral basis for it is $\lambda_1, \lambda_2, \lambda_3, \lambda_4$. Both these fields are cyclical; that is to say, the cyclical permutations

$$(\mu_1\mu_2\mu_3 \ldots \mu_8) \text{ or } (\lambda_1\lambda_2\lambda_3\lambda_4),$$

applied to any number of the field expressed in terms of the basis, gives the conjugate numbers.

Now $(\lambda_1 - \lambda_3)^2$ is unaltered by the permutation $(\lambda_1\lambda_3)(\lambda_2\lambda_4)$. Moreover it is evidently an algebraic integer. Hence

$$(\lambda_1 - \lambda_3)^2 = A_1(\lambda_1 + \lambda_3) + A_2(\lambda_2 + \lambda_4),$$

where $A_1$, $A_2$ are rational integers, so that

$$(\lambda_2 - \lambda_4)^2 = A_2(\lambda_1 + \lambda_3) + A_1(\lambda_2 + \lambda_4),$$

$$(\lambda_1 - \lambda_3)^2 + (\lambda_2 - \lambda_4)^2 = -(A_1 + A_2).$$

Now $\lambda_1$ contains, with $\omega$, its inverse $\omega^{-1}$. Hence $\lambda_1^2 - \frac{1}{4}(p-1)$ is the sum of $\frac{1}{16}(p-1)(p-5)$ primitive $p$th roots; while $\lambda_1\lambda_3$ is the sum of $\frac{1}{16}(p-1)^2$ primitive $p$th roots. Hence

$$-A_1 - A_2 = p.$$

If $\omega$ is suitably chosen it is known that

$$\lambda_1 + \lambda_3 - \lambda_2 - \lambda_4 = \sqrt{p},$$

and therefore

$$(\lambda_1 - \lambda_3)^2 = \tfrac{1}{2}p + \tfrac{1}{2}(A_1 - A_2)\sqrt{p}.$$

The algebraic integer $(\lambda_1 - \lambda_3)(\lambda_2 - \lambda_4)$ takes just two distinct values which are equal and opposite; i.e.

$$(\lambda_1 - \lambda_3)(\lambda_2 - \lambda_4) = B(\lambda_1 + \lambda_3 - \lambda_2 - \lambda_4).$$

Comparing the values of $(\lambda_1 - \lambda_3)^2(\lambda_2 - \lambda_4)^2$ derived from the last two equations

$$p^2 - (A_1 - A_2)^2 p = 4B^2 p,$$

or

$$p = (A_1 - A_2)^2 + (2B)^2.$$

Now since $p$ is congruent to unity (mod. 4), it is uniquely expressible in the form $a^2 + b^2$ where $a$ is an odd and $b$ is an even integer. Also $A_1 - A_2$ and $B$ are rational integers, so that

$$A_1 - A_2 = a,$$

where the sign of $a$ has still to be determined.

* Hilbert, *Jahresbericht der Deutschen Mathematikes-Vereinigung* (Vol. IV, p. 352), (1897).

Now from the equations

$$(\lambda_1-\lambda_3)^2=\tfrac{1}{2}p+\tfrac{1}{2}a\sqrt{p},$$

$$(\lambda_1+\lambda_3)^2=\tfrac{1}{4}(p+1)-\tfrac{1}{2}\sqrt{p},$$

there results $$-2\lambda_1\lambda_3=\tfrac{1}{8}(p-1)+\tfrac{1}{4}(a+1)\sqrt{p}.$$

Since $2\lambda_1\lambda_3$ is an algebraic integer and $\frac{1}{8}(p-1)$ is an integer, $\frac{1}{4}(a+1)\sqrt{p}$ must be an algebraic integer, and therefore

$$a\equiv -1 \pmod{4}.$$

This immediately gives the formulae

$$\lambda_1=-\tfrac{1}{4}+\tfrac{1}{4}\sqrt{p}+\tfrac{1}{4}\sqrt{2(p+a\sqrt{p})},$$

$$\lambda_2=-\tfrac{1}{4}-\tfrac{1}{4}\sqrt{p}+\tfrac{1}{4}\sqrt{2(p-a\sqrt{p})}$$

$$\lambda_3=-\tfrac{1}{4}+\tfrac{1}{4}\sqrt{p}-\tfrac{1}{4}\sqrt{2(p+a\sqrt{p})},$$

$$\lambda_4=-\tfrac{1}{4}-\tfrac{1}{4}\sqrt{p}-\tfrac{1}{4}\sqrt{2(p-a\sqrt{p})},$$

which indicate how the signs of the square roots change for the cyclical permutation $(\lambda_1\lambda_2\lambda_3\lambda_4)$.

3. The remaining step is to determine the value of $(\mu_1-\mu_5)^2$. This is an algebraic integer which is unaltered by the permutation $(\mu_1\mu_5)(\mu_2\mu_6)(\mu_3\mu_7)(\mu_4\mu_8)$, so that

$$(\mu_1-\mu_5)^2=A_1\mu_1+A_2\mu_2+A_3\mu_3+A_4\mu_4+A_1\mu_5+A_2\mu_6+A_3\mu_7+A_4\mu_8$$
$$=A_1\lambda_1+A_2\lambda_2+A_3\lambda_3+A_4\lambda_4,$$

where $A_1, A_2, A_3, A_4$ are rational integers. Using the above values of the $\lambda$'s

$$(\mu_1-\mu_5)^2=\alpha+\beta\sqrt{p}+\gamma\sqrt{2(p+a\sqrt{p})}+\delta\sqrt{2(p-a\sqrt{p})},$$

where $$\alpha=-\tfrac{1}{4}(A_1+A_2+A_3+A_4),\quad \beta=\tfrac{1}{4}(A_1-A_2+A_3-A_4),$$

$$\gamma=\tfrac{1}{4}(A_1-A_3),\qquad \delta=\tfrac{1}{4}(A_2-A_4),$$

so that $4\alpha$, $4\beta$, $4\gamma$, $4\delta$ are rational integers.

The number $(\mu_1-\mu_5)(\mu_2-\mu_6)$ is also unaltered by $(\mu_1\mu_5)(\mu_2\mu_6)(\mu_3\mu_7)(\mu_4\mu_8)$, while the sum of its four conjugate values is zero.

Hence $$(\mu_1-\mu_5)(\mu_2-\mu_6)=k_1\sqrt{p}+k_2\sqrt{2(p+a\sqrt{p})}+k_3\sqrt{2(p-a\sqrt{p})}$$

where $4k_1$, $4k_2$, $4k_3$ are rational integers. This gives

$$(\mu_3-\mu_7)(\mu_4-\mu_8)=k_1\sqrt{p}-k_2\sqrt{2(p+a\sqrt{p})}-k_3\sqrt{2(p-a\sqrt{p})},$$

so that

$$(\mu_1-\mu_5)(\mu_2-\mu_6)(\mu_3-\mu_7)(\mu_4-\mu_8)=p\,[k_1^2-2(k_2^2+k_3^2)]-2\,[a(k_2^2-k_3^2)+2bk_2k_3]\sqrt{p}.$$

Now the two conjugate values of the number on the left are equal and opposite, so that

$$k_1^2=2(k_2^2+k_3^2). \quad\text{.........................................(i)}$$

Further $$(\mu_1-\mu_5)^2(\mu_2-\mu_6)^2=[\alpha+\beta\sqrt{p}+\gamma\sqrt{2(p+a\sqrt{p})}+\delta\sqrt{2(p-a\sqrt{p})}]$$
$$\times[\alpha-\beta\sqrt{p}-\delta\sqrt{2(p+a\sqrt{p})}+\gamma\sqrt{2(p-a\sqrt{p})}],$$

while $$[(\mu_1-\mu_5)(\mu_2-\mu_6)]^2=[k_1\sqrt{p}+k_2\sqrt{2(p+a\sqrt{p})}+k_3\sqrt{2(p-a\sqrt{p})}]^2.$$

On comparing these two equal expressions, it is found they involve the following relations between the rational numbers $\alpha$, $\beta$, $\gamma$, $\delta$, $k_1$, $k_2$, $k_3$:

$$\alpha^2 - p\beta^2 = p\,[k_1^2 + 2\,(k_2^2 + k_3^2)], \qquad \text{(ii)}$$

$$b\,(\gamma^2 - \delta^2) - 2a\gamma\delta = a\,(k_2^2 - k_3^2) + 2bk_2k_3, \qquad \text{(iii)}$$

$$(\alpha + b\beta)(\gamma - \delta) - a\beta\,(\gamma + \delta) = 2k_1\,(ak_2 + bk_3), \qquad \text{(iv)}$$

$$-a\beta\,(\gamma - \delta) + (\alpha - b\beta)(\gamma + \delta) = 2k_1\,(bk_2 - ak_3), \qquad \text{(v)}$$

where $b$ is the positive square root of the $b^2$ that occurs in $p = a^2 + b^2$. The equations (iv) and (v) give

$$2pk_1k_2 = \alpha a\,(\gamma - \delta) + (\alpha b - \beta p)(\gamma + \delta),$$

$$2pk_1k_3 = (\alpha b + \beta p)(\gamma - \delta) - \alpha a\,(\gamma + \delta).$$

Entering these values of $k_2$ and $k_3$ in (iii), it becomes an identity in virtue of (i) and (ii). Entering them in (i) and (ii), it is found that $\alpha$, $\beta$, $\gamma$, $\delta$, $k_1$ must satisfy

$$\alpha^2 - p\beta^2 = 2pk_1^2 = \frac{2}{k_1^2}[(\alpha^2 + p\beta^2)(\gamma^2 + \delta^2) - 2\alpha\beta\,\{a\,(\gamma^2 - \delta^2) + 2b\gamma\delta\}]. \qquad \text{(vi)}$$

When $4\alpha$, or $\sum_1^4 (\mu_i - \mu_{i+4})^2$, is calculated in the same way that $(\lambda_1 - \lambda_3)^2 + (\lambda_2 - \lambda_4)^2$ was calculated above, it is found to be $p$ or $-p$, according as $p \equiv 1$ or $9$ (mod. 16). In either case

$$\alpha^2 = \tfrac{1}{16}p^2,$$

so that the first of equations (vi) gives

$$p = (4\beta)^2 + 2\,(4k_1)^2.$$

Now $4\beta$ and $4k_1$ are rational integers; and since $p$ is congruent to unity (mod. 8), it can be uniquely expressed in the form

$$p = a'^2 + 2b'^2,$$

where $a'$ and $b'$ are rational integers.

Hence $$\beta = \tfrac{1}{4}a', \quad k_1 = \tfrac{1}{4}b',$$

where the signs of $a'$ and $b'$ have yet to be determined.

When these values are entered in the second of equations (vi) it becomes

$$b'^4 = [4\gamma b - 4\delta\,(a \pm a')]^2 + [4\delta b + 4\gamma\,(a \mp a')]^2, \qquad \text{(vii)}$$

where the upper or lower signs are taken according as $p \equiv 1$ or $9$ (mod. 16).

Now if $c$ is a given rational number and $x$, $y$ unknown rational numbers, the equation

$$x^2 + y^2 = c^2$$

may be written $$\frac{c - x}{y} = \frac{y}{c + x} = t,$$

where $t$ is any rational number; so that its general rational solution is

$$x = c\,\frac{1 - t^2}{1 + t^2}, \quad y = c\,\frac{2t}{1 + t^2}.$$

The general rational solution of (vii) is therefore

$$4\gamma b - 4\delta (a \pm a') = b'^2 \frac{1-t^2}{1+t^2},$$

$$4\delta b + 4\gamma (a \mp a') = b'^2 \frac{2t}{1+t^2},$$

which give

$$4\gamma = \tfrac{1}{2}\left[b\frac{1-t^2}{1+t^2} + (a \pm a')\frac{2t}{1+t^2}\right],$$

$$4\delta = \tfrac{1}{2}\left[-(a \mp a')\frac{1-t^2}{1+t^2} + b\frac{2t}{1+t^2}\right].$$

Now $4\gamma$ and $4\delta$ are rational integers. Hence if the fractions $\frac{1-t^2}{1+t^2}$ and $\frac{2t}{1+t^2}$ are other than 0 and $\pm 1$ (i.e. if they have an effective denominator $d$), then $d$ must be a factor of $b'^2$ from the first form of the equations and of $a'$ from the second form. Hence, since $a'$ and $b'$ are relatively prime, the only values of $\frac{1-t^2}{1+t^2}$ and $\frac{2t}{1+t^2}$ are 0 and $\pm 1$. It follows that there are only four possible pairs of values for $\gamma$ and $\delta$, viz.

$$\gamma = \tfrac{1}{8}(a \pm a'),\quad \gamma = -\tfrac{1}{8}(a \pm a'),\quad \gamma = \tfrac{1}{8}b,\quad \gamma = -\tfrac{1}{8}b,$$
$$\delta = \tfrac{1}{8}b,\quad \delta = -\tfrac{1}{8}b,\quad \delta = -\tfrac{1}{8}(a \mp a'),\quad \delta = \tfrac{1}{8}(a \mp a').$$

These possible values of $\gamma$ and $\delta$, and the sign of $a'$ may be dealt with as follows. The equations

$$(\mu_1 - \mu_5)^2 + (\mu_3 - \mu_7)^2 = \pm \tfrac{1}{2}p + \tfrac{1}{2}a'\sqrt{p},$$

$$(\mu_1 + \mu_5)^2 + (\mu_3 + \mu_7)^2 = \tfrac{1}{8}(\sqrt{p} - 1)^2 + \tfrac{1}{4}(p + a\sqrt{p})$$

give

$$-2(\mu_1\mu_5 + \mu_3\mu_7) = \tfrac{1}{16}(p-1) + \tfrac{1}{4}[a' - \tfrac{1}{2}(a-1)]\sqrt{p}, \qquad p \equiv 1 \pmod{16},$$
$$= -\tfrac{1}{16}(7p+1) + \tfrac{1}{4}[a' - \tfrac{1}{2}(a-1)]\sqrt{p}, \quad p \equiv 9 \pmod{16},$$

so that in either case

$$a' \equiv \tfrac{1}{2}(a-1), \pmod{4}.$$

The equations

$$(\mu_1 - \mu_5)^2 - (\mu_3 - \mu_7)^2 = 2\gamma\sqrt{2(p + a\sqrt{p})} + 2\delta\sqrt{2(p - a\sqrt{p})},$$

$$(\mu_1 + \mu_5)^2 - (\mu_3 + \mu_7)^2 = \tfrac{1}{4}(\sqrt{p} - 1)\sqrt{2(p + a\sqrt{p})}$$
$$= \tfrac{1}{4}(a-1)\sqrt{2(p + a\sqrt{p})} + \tfrac{1}{4}b\sqrt{2(p - a\sqrt{p})}$$

give

$$\mu_3\mu_7 - \mu_1\mu_5 = \tfrac{1}{16}(8\gamma - a + 1)\sqrt{2(p + a\sqrt{p})} + \tfrac{1}{16}(8\delta - b)\sqrt{2(p - a\sqrt{p})}$$
$$= \tfrac{1}{8}(8\gamma - a + 1)(\lambda_1 + \lambda_3) + \tfrac{1}{8}(8\delta - b)(\lambda_2 + \lambda_4). \quad \text{..........................(viii)}$$

Since $\mu_3\mu_7 - \mu_1\mu_5$ is an algebraic integer, $\tfrac{1}{8}(8\gamma - a + 1)$ and $\tfrac{1}{8}(8\delta - b)$ must be rational integers. For the last two possible pairs of values of $\gamma$ and $\delta$, $\tfrac{1}{8}(8\gamma - a + 1)$ is $\tfrac{1}{8}(1 - a \pm b)$. Since $a \equiv -1$, $b \equiv 0 \pmod{4}$, this number cannot be an integer. Hence only the first two pairs of values of $\gamma$ and $\delta$ can occur.

If $p \equiv 1 \pmod{16}$, $a^2$ is of the form $(8m \pm 1)^2$; and since $a \equiv -1 \pmod{4}$, in this case $a \equiv -1 \pmod{8}$. Hence $a' \equiv \tfrac{1}{2}(a-1) \pmod{4}$, $\equiv -1 \pmod{4}$, $= -1$ or $3 \pmod{8}$.

If $a' \equiv -1 \pmod{8}$, $8\gamma = a + a'$, $\quad 8\delta = b$, ⎫
$\quad a' \equiv 3 \pmod{8}$, $\quad 8\gamma = -a - a'$, $8\delta = -b$ ⎭ make the coefficients in (viii) integral.

When $p \equiv 9$ (mod. 16), $a^2$ is of the form $(8m \pm 3)^2$; and since $a \equiv -1$ (mod. 4), in this case $a \equiv 3$ (mod. 8). Hence $a' \equiv \frac{1}{2}(a-1)$ (mod. 4), $\equiv 1$ (mod. 4), $\equiv 1$ or $-3$ (mod. 8).

$$\left.\begin{array}{ll} \text{If } a' \equiv 1 \text{ (mod. 8)}, & 8\gamma = a - a', \quad 8\delta = b, \\ a' \equiv -3 \text{ (mod. 8)}, & 8\gamma = -a + a', \quad 8\delta = -b \end{array}\right\} \text{make the coefficients in (viii) integral.}$$

To sum up, the results may be expressed as follows.

If
$$p = a^2 + b^2 = a'^2 + 2b'^2,$$
where $p$ is a prime congruent to unity (mod. 8), while $a$, $b$ and $a'$ are completely defined by
$$a \equiv -1 \text{ (mod. 4)}, \quad b > 0, \quad a' \equiv \tfrac{1}{2}(a-1) \text{ (mod. 4)},$$
then when
$$p \equiv 1 \text{ (mod. 16)}, \quad a' \equiv -1 \text{ (mod. 8)},$$
$$4(\mu_1 - \mu_5)^2 = p + a'\sqrt{p} + \tfrac{1}{2}(a+a')\sqrt{2(p + a\sqrt{p})} + \tfrac{1}{2}b\sqrt{2(p - a\sqrt{p})},$$
when
$$p \equiv 1 \text{ (mod. 16)}, \quad a' \equiv 3 \text{ (mod. 8)},$$
$$4(\mu_1 - \mu_5)^2 = p + a'\sqrt{p} - \tfrac{1}{2}(a+a')\sqrt{2(p + a\sqrt{p})} - \tfrac{1}{2}b\sqrt{2(p - a\sqrt{p})},$$
when
$$p \equiv 9 \text{ (mod. 16)}, \quad a' \equiv 1 \text{ (mod. 8)},$$
$$4(\mu_1 - \mu_5)^2 = -p + a'\sqrt{p} + \tfrac{1}{2}(a-a')\sqrt{2(p + a\sqrt{p})} + \tfrac{1}{2}b\sqrt{2(p - a\sqrt{p})},$$
when
$$p \equiv 9 \text{ (mod. 16)}, \quad a' \equiv -3 \text{ (mod. 8)},$$
$$4(\mu_1 - \mu_5)^2 = -p + a'\sqrt{p} - \tfrac{1}{2}(a-a')\sqrt{2(p + a\sqrt{p})} - \tfrac{1}{2}b\sqrt{2(p - a\sqrt{p})},$$
these possibilities covering all cases, while in each case
$$4(\mu_1 + \mu_5) = -1 + \sqrt{p} + \sqrt{2(p + a\sqrt{p})}.$$

4. When the values of $\alpha, \beta, \gamma, \delta$ that have now been determined are entered in the equations
$$2pk_1k_2 = \alpha a(\gamma - \delta) + (\alpha b - \beta p)(\gamma + \delta),$$
$$2pk_1k_3 = (\alpha b + \beta p)(\gamma - \delta) - \alpha a(\gamma + \delta),$$
they give
$$\begin{array}{llll} p \equiv 1 \text{ (mod. 16)}, & a' \equiv -1 \text{ (mod. 8)}, & 32k_1k_2 = b'^2, & 32k_1k_3 = -b'^2, \\ p \equiv 1 \text{ (mod. 16)}, & a' \equiv 3 \text{ (mod. 8)}, & 32k_1k_2 = -b'^2, & 32k_1k_3 = b'^2, \\ p \equiv 9 \text{ (mod. 16)}, & a' \equiv 1 \text{ (mod. 8)}, & 32k_1k_2 = -b'^2, & 32k_1k_3 = b'^2, \\ p \equiv 9 \text{ (mod. 16)}, & a' \equiv -3 \text{ (mod. 8)}, & 32k_1k_2 = b'^2, & 32k_1k_3 = -b'^2. \end{array}$$

When $p \equiv 1$ (mod. 16), $(\mu_1 - \mu_5)^2$ is a real positive number so that, with the specification that has been given of the symbol $\sqrt{(\mu_1 - \mu_5)^2}$, it follows that $(\mu_1 - \mu_5)(\mu_2 - \mu_6) + (\mu_3 - \mu_7)(\mu_4 - \mu_8)$, or $2k_1\sqrt{p}$, is a real positive number. Hence, if $b' > 0$,
$$p \equiv 1 \text{ (mod. 16)}, \quad a' \equiv -1 \text{ (mod. 8)}, \quad k_1 = \tfrac{1}{4}b', \quad k_2 = -k_3 = \tfrac{1}{8}b',$$
$$p \equiv 1 \text{ (mod. 16)}, \quad a' \equiv 3 \text{ (mod. 8)}, \quad k_1 = \tfrac{1}{4}b', \quad k_2 = -k_3 = -\tfrac{1}{8}b'.$$

When $p \equiv 9$ (mod. 16), $(\mu_1 - \mu_5)^2$ is a real negative number, and it is necessary to give a specification of the symbol $\sqrt{(\mu_1 - \mu_5)^2}$. If it is specified as that square root in which the coefficient of $i$ is positive, it follows that $(\mu_1 - \mu_5)(\mu_2 - \mu_6) + (\mu_3 - \mu_7)(\mu_4 - \mu_8)$, or $2k_1\sqrt{p}$, is a real negative number.

Hence, if $b' > 0$,

$$p \equiv 9 \text{ (mod. 16)}, \quad a' \equiv \phantom{-}1 \text{ (mod. 8)}, \quad k_1 = -\tfrac{1}{4}b', \quad k_2 = -k_3 = \tfrac{1}{8}b',$$

$$p \equiv 9 \text{ (mod. 16)}, \quad a' \equiv -3 \text{ (mod. 8)}, \quad k_1 = -\tfrac{1}{4}b', \quad k_2 = -k_3 = -\tfrac{1}{8}b'.$$

Denoting the four cases, as regards the values of $p$ and $a'$, in the above order, by (i), (ii), (iii), (iv), these results are equivalent to the formulae

$$\text{(i)} \quad (\mu_1 - \mu_5)(\mu_2 - \mu_6) = \tfrac{1}{2}b'(\lambda_1 - \lambda_2),$$

$$\text{(ii)} \quad (\mu_1 - \mu_5)(\mu_2 - \mu_6) = \tfrac{1}{2}b'(\lambda_3 - \lambda_4),$$

$$\text{(iii)} \quad (\mu_1 - \mu_5)(\mu_2 - \mu_6) = -\tfrac{1}{2}b'(\lambda_3 - \lambda_4),$$

$$\text{(iv)} \quad (\mu_1 - \mu_5)(\mu_2 - \mu_6) = -\tfrac{1}{2}b'(\lambda_1 - \lambda_2).$$

5. To complete the formulæ for the multiplication of the differences $\mu_i - \mu_{i+4}$, it is necessary to calculate directly that for $(\mu_1 - \mu_5)(\mu_3 - \mu_7)$, as it cannot be obtained by cyclical interchange from those just given.

In cases (i) and (ii)

$$(\mu_1 - \mu_5)(\mu_3 - \mu_7) = \sqrt{(\alpha + \beta\sqrt{p})^2 - [\gamma\sqrt{2(p + a\sqrt{p})} + \delta\sqrt{2(p - a\sqrt{p})}]^2}$$

$$= \tfrac{1}{4}\sqrt{p(a'^2 + b'^2 - aa') + (pa' - aa'^2 - ab'^2)\sqrt{p}}$$

$$= \tfrac{1}{8}\sqrt{[b\sqrt{2(p + a\sqrt{p})} - (a - a')\sqrt{2(p - a\sqrt{p})}]^2}.$$

Now in these cases $(\mu_1 - \mu_5)(\mu_3 - \mu_7)$ is positive. Hence

$$(\mu_1 - \mu_5)(\mu_3 - \mu_7) = \tfrac{1}{8}b\sqrt{2(p + a\sqrt{p})} - \tfrac{1}{8}(a - a')\sqrt{2(p - a\sqrt{p})}$$

$$= \tfrac{1}{4}b(\lambda_1 - \lambda_3) - \tfrac{1}{4}(a - a')(\lambda_2 - \lambda_4).$$

Similarly in cases (iii) and (iv), when $(\mu_1 - \mu_5)(\mu_3 - \mu_7)$ is negative,

$$(\mu_1 - \mu_5)(\mu_3 - \mu_7) = \tfrac{1}{4}\sqrt{p(a'^2 + b'^2 + aa') - (pa' + aa'^2 + bb'^2)\sqrt{p}}$$

$$= \tfrac{1}{8}\sqrt{[b\sqrt{2(p + a\sqrt{p})} - (a + a')\sqrt{2(p - a\sqrt{p})}]^2}$$

$$= -\tfrac{1}{8}b\sqrt{2(p + a\sqrt{p})} + \tfrac{1}{8}(a + a')\sqrt{2(p - a\sqrt{p})}$$

$$= -\tfrac{1}{4}b(\lambda_1 - \lambda_3) + \tfrac{1}{4}(a + a')(\lambda_2 - \lambda_4).$$

# XXI. *Congruences with respect to Composite Moduli.*

By Major P. A. MacMahon, R.A., F.R.S.

[*Received* 16 April, *Read* 3 May, 1920.]

The object of the present communication is to put together certain results in the Theory of the Residues of Powers with respect to composite moduli which seem to be worthy of preservation. The generalised Fermat Theorem states that if $M$ be any integer, the congruence

$$x^{\phi(M)} \equiv 1 \bmod M$$

has $\phi(M)$ incongruent roots which are the $\phi(M)$ numbers less than and prime to $M$. If $M$ have either of the forms (i) a power of an uneven prime, (ii) twice the power of an uneven prime, (iii) the number 4, the congruence has roots which appertain to the exponent $\phi(M)$ and then are known as primitive roots of the congruence of the number $M$. In no other case does the congruence possess roots which appertain to the exponent $\phi(M)$. Further if $\delta$ be a division of $\phi(M)$ the congruence

$$x^{\delta} \equiv 1 \bmod M$$

has $\delta$ roots, $\phi(\delta)$ of which appertain to the exponent $\delta$ whenever $M$ has one of the three forms specified.

1. In order to regard the matter from a general point of view it is convenient to separate the assemblage of all integers into an infinite number of categories.

Denoting uneven primes by

$$p_1, p_2, p_3, \ldots,$$

I define the $s$th category as including numbers of the four forms

$$p_1^{\pi_1} p_2^{\pi_2} \ldots p_s^{\pi_s},$$

$$2p_1^{\pi_1} p_2^{\pi_2} \ldots p_s^{\pi_s},$$

$$2^2 p_1^{\pi_1} p_2^{\pi_2} \ldots p_{s-1}^{\pi_{s-1}},$$

$$2^{\alpha} p_1^{\pi_1} p_2^{\pi_2} \ldots p_{s-2}^{\pi_{s-2}}, \text{ where } \alpha > 2,$$

where observe that putting $s=1$, the first category involves (exceptionally) numbers of the three forms

$$p_1^{\pi_1},\ 2p_1^{\pi_1},\ 2^2,$$

the fourth form not existing and that it comprises precisely and exclusively those numbers which possess primitive roots.

For the first category Tables exist which shew up to a certain value of $M$ the roots which appertain to every divisor of $\phi(M)$*.

* *Die unbestimmte Analytik* by Hermann Scheffler gives such a Table for all prime moduli which are $<100$ and other more extensive results have been given in *Crelle* and elsewhere.

In respect of moduli which are included in categories other than the first it is well known* that the highest exponent to which roots appertain is if

$$M = 2^{\alpha} p_1^{\pi_1} p_2^{\pi_2} \dots p_s^{\pi_s}$$

the least common multiple of

$$\phi(2^{\alpha}),\ \phi(p_1^{\pi_1}),\ \phi(p_2^{\pi_2}), \dots \phi(p_s^{\pi_s}),$$

if $\alpha \leqslant 2$, and the least common multiple of

$$\phi^2(2^{\alpha}),\ \phi(p_1^{\pi_1}),\ \phi(p_2^{\pi_2}), \dots \phi(p_s^{\pi_s}),$$

if $\alpha > 2$.

I denote this number by

$$\theta(M),$$

where observe that

$$\theta(M) = \phi(M) \text{ or } \leqslant \tfrac{1}{2}\phi(M)$$

according as $M$ is or is not in the first category.

For numbers in all categories we may state:

"The congruence

$$x^{\theta(M)} \equiv 1 \mod M$$

possesses $\phi(M)$ incongruent roots which are the $\phi(M)$ numbers less than and prime to $M$."

It is clear that $\theta(M)$ is a divisor of $\phi(M)$.

2. It has been established by previous writers that, if $M_s$ be a number included in the $s$th category, the congruence

$$x^2 \equiv 1 \mod M_s$$

possesses $2^s$ roots exactly and that of these, $2^s - 1$ appertain to the exponent 2.

We may in fact so define the $s$th category, simply saying that it includes all numbers $M_s$ which have the property that the congruence

$$x^2 \equiv 1 \mod M_s$$

possesses exactly $2^s$ roots.

The congruence

$$x^{\theta(M)} \equiv 1 \mod M$$

has roots which appertain to the exponent $\theta(M)$. They are primitive roots of the congruence.

Further if $\delta$ be a divisor of $\theta(M)$ the roots which appertain to the exponent $\delta$ may be termed primitive roots of the congruence

$$x^{\delta} \equiv 1 \mod M.$$

When $M$ belongs to the first category this congruence possesses $\phi(\delta)$ primitive roots but in the cases of other categories this is not so. It is in fact easy to verify for the special case $\delta = 2$, that if $\alpha_2$ be one primitive root, the remaining $2^s - 2$ primitive roots are not all congruent to powers of $\alpha_2$.

* Serret, *Cours d'Algèbre Supérieure*, 5th Ed. 2nd Vol. pp. 50 et seq.

So also in the case of the exponent $\delta$, any one primitive root $g$ will certainly give rise, through the series of powers

$$g, g^2, g^3, \dots g^\delta, \text{ where } g^\delta \equiv 1,$$

to $\phi(\delta)$ primitive roots, because the residue of $g^k$ is a primitive root if $k$ be less than and prime to $\delta$; but in general there are other primitive roots. All that we can assert at present is that if $N_\delta$ denote the number of primitive roots of the congruence

$$x^\delta \equiv 1 \bmod M,$$

$\delta$ being a divisor of $\theta(M)$,

$$N_\delta \equiv 0 \bmod \phi(\delta).$$

We write therefore

$$N_\delta = S_\delta \phi(\delta),$$

so that the $N_\delta$ roots occur in $S_\delta$ periods or sets of power-residues.

$S_\delta$ is an arithmetic quantity—an integer—whose value has not yet been determined.

Of course for the first category $S_\delta$ is invariably unity.

If $\delta_1, \delta_2$ be relatively prime divisors of $\theta(M)$ it is known that

$$N_{\delta_1} N_{\delta_2} = N_{\delta_1\delta_2},$$

and since

$$\phi(\delta_1)\phi(\delta_2) = \phi(\delta_1\delta_2)$$

we find that

$$S_{\delta_1} S_{\delta_2} = S_{\delta_1\delta_2}.$$

Denote by

$$P_\delta$$

the number of roots of the congruence

$$x^\delta \equiv 1 \bmod M,$$

so that, in particular,

$$P_{\theta(M)} = \phi(M).$$

Also if $\delta'$ be a divisor of $\delta$

$$\Sigma N_{\delta'} = P_\delta.$$

In the first place we evaluate

$$P_\delta.$$

3. Let

$$M = 2^\alpha p_1^{\pi_1} p_2^{\pi_2} \dots p_s^{\pi_s},$$

and denote the highest common divisor of the integers $a, b$ by

$$\{a, b\}.$$

It follows from well-known principles* that the congruence

$$x^\delta \equiv 1 \bmod 2^\alpha$$

possesses

$$b\{\delta. \phi^2(2^\alpha)\} \text{ roots},$$

where $b$ is 1 or 2 according as $\delta$ is uneven or even.

Also the congruence

$$x^\delta \equiv 1 \bmod p_1^{\pi_1}$$

possesses

$$\{\delta, \phi(p_1^{\pi_1})\} \text{ roots}.$$

Hence, if $\delta$ be a divisor of $\theta(M)$, the number of roots of the congruence

$$x^\delta \equiv 1 \bmod M$$

is given by

$$P_\delta = b\{\delta, \phi^2(2^\alpha)\}, \{\delta, \phi(p_1^{\pi_1})\}, \{\delta, \phi(p_2^{\pi_2})\}, \dots \{\delta, \phi(p_s^{\pi_s})\},$$

wherein, if $M$ be uneven, the factor

$$b\{\delta, \phi^2(2^\alpha)\}$$

must be deleted.

* Serret, Vol. II, p. 89.

Observe that since

$$\phi(p^{\pi}) \text{ is even}$$

$$P_2 = 2^s$$

for moduli of each of the forms of the $s$th category, in agreement with the known result.

If $\delta$ be a power of an uneven prime $p$, each factor in $\{\ \}$ is either unity or a power of $p$. This shews that

$$P_{p^{\pi}} \text{ is equal to a power of } p.$$

It is well known that, if $\delta_1$, $\delta_2$ be relatively prime,

$$P_{\delta_1} P_{\delta_2} = P_{\delta_1 \delta_2}.$$

4. We can now evaluate

$$N_{r_1^{\rho_1} r_2^{\rho_2} r_3^{\rho_3} \ldots},$$

where $r_1^{\rho_1} r_2^{\rho_2} r_3^{\rho_3} \ldots$ is a divisor of $\theta(M)$ and of $\phi(M)$.

For $$P_{p^{\pi}} = N_1 + N_p + N_{p^2} + \ldots + N_{p^{\pi}},$$

so that $$N_{p^{\pi}} = P_{p^{\pi}} - P_{p^{\pi-1}};$$

leading to the formula

$$N_{r_1^{\rho_1} r_2^{\rho_2} r_3^{\rho_3} \ldots} = (P_{r_1^{\rho_1}} - P_{r_1^{\rho_1 - 1}})(P_{r_2^{\rho_2}} - P_{r_2^{\rho_2 - 1}})(P_{r_3^{\rho_3}} - P_{r_3^{\rho_3 - 1}}) \ldots,$$

where

$$P_{r^{\rho}} = b\{r^{\rho}, \phi^2(2^{\alpha})\}, \{r^{\rho}, \phi(p_1^{\pi_1})\}, \{r^{\rho}, \phi(p_2^{\pi_2})\}, \ldots \{r^{\rho}, \phi(p_s^{\pi_s})\}.$$

Thence

$$S_{r_1^{\rho_1} r_2^{\rho_2} r_3^{\rho_3} \ldots} = \frac{(P_{r_1^{\rho_1}} - P_{r_1^{\rho_1 - 1}})(P_{r_2^{\rho_2}} - P_{r_2^{\rho_2 - 1}})(P_{r_3^{\rho_3}} - P_{r_3^{\rho_3 - 1}}) \ldots}{\phi(r_1^{\rho_1} r_2^{\rho_2} r_3^{\rho_3} \ldots)},$$

the general expression for the number of periods in which the

$$N_{r_1^{\rho_1} r_2^{\rho_2} r_3^{\rho_3} \ldots} \text{ numbers}$$

occur.

Observe that we may write

$$S_{r_1^{\rho_1} r_2^{\rho_2} r_3^{\rho_3} \ldots} = \Pi \frac{P_{r_1^{\rho_1}} - P_{r_1^{\rho_1 - 1}}}{r_1^{\rho_1} - r_1^{\rho_1 - 1}}.$$

The annexed Table embraces all moduli of the second and third categories up to the points where moduli of the fourth category begin to appear. The smallest number belonging to the fourth category is $120 = 2^3 . 3 . 5$.

The two categories are given separately and in a form to facilitate the verification of the theorems given in the paper.

# TABLES OF PRIMITIVE ROOTS OF CONGRUENCES

## THE SECOND CATEGORY

| Modulus $M$ | $\theta(M)$ | Exponent $\delta$ | Number of Roots $P_\delta$ | $\phi(\delta)$ | Number of Periods $S_\delta$ | Number of Primitive Roots $N_\delta$ | Primitive Roots of Congruence $x^\delta \equiv 1$ |
|---|---|---|---|---|---|---|---|
| $2^3$ | 2 | 1 | 1 | 1 | 1 | 1 | 1 |
| | | 2 | 4 | 1 | 3 | 3 | 3, 5, 7 |
| $2^2.3$ | 2 | 1 | 1 | 1 | 1 | 1 | 1 |
| | | 2 | 4 | 1 | 3 | 3 | 5, 7, 11 |
| 3.5 | 4 | 1 | 1 | 1 | 1 | 1 | |
| | | 2 | 4 | 1 | 3 | 3 | 4, 11, 14 |
| | | 4 | 8 | 2 | 2 | 4 | 2, 7, 8, 13 |
| $2^4$ | 4 | 1 | 1 | 1 | 1 | 1 | 1 |
| | | 2 | 4 | 1 | 3 | 3 | 7, 9, 15 |
| | | 4 | 8 | 2 | 2 | 4 | 3, 5, 11, 13 |
| $2^2.5$ | 4 | 1 | 1 | 1 | 1 | 1 | 1 |
| | | 2 | 4 | 1 | 3 | 3 | 9, 11, 19 |
| | | 4 | 8 | 2 | 2 | 4 | 3, 7, 13, 17 |
| 3.7 | 6 | 1 | 1 | 1 | 1 | 1 | 1 |
| | | 2 | 4 | 1 | 3 | 3 | 8, 13, 20 |
| | | 3 | 3 | 2 | 1 | 2 | 4, 16 |
| | | 6 | 12 | 2 | 3 | 6 | 2, 5, 10, 11, 17, 19 |
| $2^2.7$ | 6 | 1 | 1 | 1 | 1 | 1 | 1 |
| | | 2 | 4 | 1 | 3 | 3 | 13, 15, 27 |
| | | 3 | 3 | 2 | 1 | 2 | 9, 25 |
| | | 6 | 12 | 2 | 3 | 6 | 3, 5, 11, 17, 19, 23 |
| 2.3.5 | 4 | 1 | 1 | 1 | 1 | 1 | 1 |
| | | 2 | 4 | 1 | 3 | 3 | 11, 19, 29 |
| | | 4 | 8 | 2 | 2 | 4 | 7, 13, 17, 23 |
| $2^5$ | 8 | 1 | 1 | 1 | 1 | 1 | 1 |
| | | 2 | 4 | 1 | 3 | 3 | 15, 17, 31 |
| | | 4 | 8 | 2 | 2 | 4 | 7, 9, 23, 25 |
| | | 8 | 16 | 4 | 2 | 8 | 3, 5, 11, 13, 19, 21, 27, 29 |
| 3.11 | 10 | 1 | 1 | 1 | 1 | 1 | 1 |
| | | 2 | 4 | 1 | 3 | 3 | 10, 23, 32 |
| | | 5 | 5 | 4 | 1 | 4 | 4, 16, 25, 31 |
| | | 10 | 20 | 4 | 3 | 12 | 2, 5, 7, 8, 13, 14, 17, 19, 20, 26, 28, 29 |
| 5.7 | 12 | 1 | 1 | 1 | 1 | 1 | 1 |
| | | 2 | 4 | 1 | 3 | 3 | 6, 29, 34 |
| | | 3 | 3 | 2 | 1 | 2 | 11, 16 |
| | | 4 | 8 | 2 | 2 | 4 | 8, 13, 22, 27 |
| | | 6 | 12 | 2 | 3 | 6 | 4, 9, 19, 24, 26, 31 |
| | | 12 | 24 | 4 | 2 | 8 | 2, 3, 12, 17, 18, 23, 32, 33 |
| $2^2.3^2$ | 6 | 1 | 1 | 1 | 1 | 1 | 1 |
| | | 2 | 4 | 1 | 3 | 3 | 17, 19, 35 |
| | | 3 | 3 | 2 | 1 | 2 | 13, 25 |
| | | 6 | 12 | 2 | 3 | 6 | 5, 7, 11, 23, 29, 31 |

| Modulus $M$ | $\theta(M)$ | Exponent $\delta$ | Number of Roots $P_\delta$ | $\phi(\delta)$ | Number of Periods $S_\delta$ | Number of Primitive Roots $N_\delta$ | Primitive Roots of Congruence $x^\delta \equiv 1$ |
|---|---|---|---|---|---|---|---|
| 3.13 | 12 | 1 | 1 | 1 | 1 | 1 | 1 |
| | | 2 | 4 | 1 | 3 | 3 | 14, 25, 38 |
| | | 3 | 3 | 2 | 1 | 2 | 16, 22 |
| | | 4 | 8 | 2 | 2 | 4 | 5, 8, 31, 34 |
| | | 6 | 12 | 2 | 3 | 6 | 4, 10, 17, 23, 29, 35 |
| | | 12 | 24 | 4 | 2 | 8 | 2, 7, 11, 19, 20, 28, 32, 37 |
| 3.7 | 6 | 1 | 1 | 1 | 1 | 1 | 1 |
| | | 2 | 4 | 1 | 3 | 3 | 12, 29, 41 |
| | | 3 | 3 | 2 | 1 | 2 | 25, 37 |
| | | 6 | 12 | 2 | 3 | 6 | 5, 7, 11, 23, 29, 31 |
| $2^2$.11 | 10 | 1 | 1 | 1 | 1 | 1 | 1 |
| | | 2 | 4 | 1 | 3 | 3 | 21, 23, 43 |
| | | 5 | 5 | 4 | 1 | 4 | 5, 9, 25, 37 |
| | | 10 | 20 | 4 | 3 | 12 | 3, 7, 13, 15, 17, 19, 27, 29, 31, 35, 39, 41 |
| $3^2$.5 | 12 | 1 | 1 | 1 | 1 | 1 | 1 |
| | | 2 | 4 | 1 | 3 | 3 | 19, 26, 44 |
| | | 3 | 3 | 2 | 1 | 2 | 16, 31 |
| | | 4 | 8 | 2 | 2 | 4 | 8, 17, 28, 37 |
| | | 6 | 12 | 2 | 3 | 6 | 4, 11, 14, 29, 34, 41 |
| | | 12 | 24 | 4 | 2 | 8 | 2, 7, 13, 22, 23, 32, 38, 43 |
| 3.17 | 16 | 1 | 1 | 1 | 1 | 1 | 1 |
| | | 2 | 4 | 1 | 3 | 3 | 16, 35, 50 |
| | | 4 | 8 | 2 | 2 | 4 | 4, 13, 38, 47 |
| | | 8 | 16 | 4 | 2 | 8 | 2, 8, 19, 25, 26, 32, 43, 49 |
| | | 16 | 32 | 8 | 2 | 16 | 5, 7, 10, 11, 14, 20, 22, 23, 28, 29, 31, 37, 40, 41, 44, 46 |
| 5.11 | 20 | 1 | 1 | 1 | 1 | 1 | 1 |
| | | 2 | 4 | 1 | 3 | 3 | 21, 34, 54 |
| | | 4 | 8 | 2 | 2 | 4 | 12, 23, 32, 43 |
| | | 5 | 5 | 4 | 1 | 4 | 16, 26, 31, 36 |
| | | 10 | 20 | 4 | 3 | 12 | 4, 6, 9, 14, 19, 24, 29, 39, 41, 46, 49, 51 |
| | | 20 | 40 | 8 | 2 | 16 | 2, 3, 7, 8, 13, 17, 18, 27, 28, 37, 38, 42, 47, 48, 52, 53 |
| $2^2$.13 | 12 | 1 | 1 | 1 | 1 | 1 | 1 |
| | | 2 | 4 | 1 | 3 | 3 | 25, 27, 51 |
| | | 3 | 3 | 2 | 1 | 2 | 9, 29 |
| | | 4 | 8 | 2 | 2 | 4 | 5, 21, 31, 47 |
| | | 6 | 12 | 2 | 3 | 6 | 3, 17, 23, 35, 43, 49 |
| | | 12 | 24 | 4 | 2 | 8 | 7, 11, 15, 19, 33, 37, 41, 45 |
| 3.19 | 18 | 1 | 1 | 1 | 1 | 1 | 1 |
| | | 2 | 4 | 1 | 3 | 3 | 20, 37, 56 |
| | | 3 | 3 | 2 | 1 | 2 | 7, 49 |
| | | 6 | 12 | 2 | 3 | 6 | 8, 11, 26, 31, 46, 50 |
| | | 9 | 9 | 6 | 1 | 6 | 4, 16, 25, 28, 43, 55 |
| | | 18 | 36 | 6 | 3 | 18 | 2, 5, 10, 13, 14, 17, 22, 23, 29, 32, 34, 35, 40, 41 44, 47, 52, 53 |

| Modulus $M$ | $\theta(M)$ | Exponent $\delta$ | Number of Roots $P_\delta$ | $\phi(\delta)$ | Number of Periods $S_\delta$ | Number of Primitive Roots $N_\delta$ | Primitive Roots of Congruence $x^\delta \equiv 1$ |
|---|---|---|---|---|---|---|---|
| $3^2.7$ | 6 | 1 | 1 | 1 | 1 | 1 | 1 |
| | | 2 | 4 | 1 | 3 | 3 | 8, 55, 62 |
| | | 3 | 9 | 2 | 4 | 8 | 4, 16, 22, 25, 37, 43, 46, 58 |
| | | 6 | 36 | 2 | 12 | 24 | 2, 5, 10, 11, 13, 17, 19, 20, 23, 26, 29, 31, 32, 34, 38, 40, 41, 44, 47, 50, 52, 53, 59, 61 |
| $2^6$ | 16 | 1 | 1 | 1 | 1 | 1 | 1 |
| | | 2 | 4 | 1 | 3 | 3 | 31, 33, 63 |
| | | 4 | 8 | 2 | 2 | 4 | 15, 17, 47, 49 |
| | | 8 | 16 | 4 | 2 | 8 | 7, 9, 23, 25, 39, 41, 55, 57 |
| | | 16 | 32 | 8 | 2 | 16 | 3, 5, 11, 13, 19, 21, 27, 29, 35, 37, 43, 45, 51, 53, 59, 61 |
| 5.13 | 12 | 1 | 1 | 1 | 1 | 1 | 1 |
| | | 2 | 4 | 1 | 3 | 3 | 14, 51, 64 |
| | | 3 | 3 | 2 | 1 | 2 | 16, 61 |
| | | 4 | 16 | 2 | 6 | 12 | 8, 12, 18, 21, 27, 31, 34, 38, 44, 47, 53, 57 |
| | | 6 | 12 | 2 | 3 | 6 | 4, 9, 29. 36, 49, 56 |
| | | 12 | 48 | 4 | 6 | 24 | 2, 3, 6, 7, 11, 17 19, 22, 23, 24, 28, 32, 33, 37, 41, 42, 43, 46, 48, 54, 58, 59, 62, 63 |
| 2.3.11 | 10 | 1 | 1 | 1 | 1 | 1 | 1 |
| | | 2 | 4 | 1 | 3 | 3 | 23, 45, 65 |
| | | 5 | 5 | 4 | 1 | 4 | 25, 31, 37, 49 |
| | | 10 | 20 | 4 | 3 | 12 | 5, 7, 13, 17, 19, 29, 35, 41, 47, 53, 59, 61 |
| $2^2.17$ | 16 | 1 | 1 | 1 | 1 | 1 | 1 |
| | | 2 | 4 | 1 | 3 | 3 | 33, 35, 67 |
| | | 4 | 8 | 2 | 2 | 4 | 13, 21, 47, 55 |
| | | 8 | 16 | 4 | 2 | 8 | 9, 15, 19, 25, 43, 49, 53, 59 |
| | | 16 | 32 | 8 | 2 | 16 | 3, 5, 7, 11, 23, 27, 29, 31, 37, 39, 41, 45, 57, 61, 63, 65 |
| 3.23 | 22 | 1 | 1 | 1 | 1 | 1 | 1 |
| | | 2 | 4 | 1 | 3 | 3 | 22, 47, 68 |
| | | 11 | 11 | 10 | 1 | 10 | 4, 13, 16, 25, 31, 49, 52, 55, 58, 64 |
| | | 22 | 44 | 10 | 3 | 30 | 2, 5, 7, 8, 10, 11, 14, 17, 19, 20, 26, 28, 29, 32, 34, 35, 37, 38, 40, 41, 43, 44, 50, 53, 56, 59, 61, 62, 65, 67 |
| 2.5.7 | 12 | 1 | 1 | 1 | 1 | 1 | 1 |
| | | 2 | 4 | 1 | 3 | 3 | 29, 41, 69 |
| | | 3 | 3 | 2 | 1 | 2 | 11, 51 |
| | | 4 | 8 | 2 | 2 | 4 | 13, 27, 43, 57 |
| | | 6 | 12 | 2 | 3 | 6 | 9, 19, 31, 39, 59, 61 |
| | | 12 | 24 | 4 | 2 | 8 | 3, 17, 23, 33, 37, 47, 53, 67 |
| $3.5^2$ | 20 | 1 | 1 | 1 | 1 | 1 | 1 |
| | | 2 | 4 | 1 | 3 | 3 | 26, 49, 74 |
| | | 4 | 8 | 2 | 2 | 4 | 7, 32, 43, 68 |
| | | 5 | 5 | 4 | 1 | 4 | 16, 31, 46, 61 |
| | | 10 | 20 | 4 | 3 | 12 | 4, 11, 14, 19, 29, 34, 41, 44, 56, 59, 64, 71 |
| | | 20 | 40 | 8 | 2 | 16 | 2, 8, 13, 17, 22, 23, 28, 37, 38, 47, 52, 53, 58, 62, 67, 73 |

| Modulus $M$ | $\theta(M)$ | Exponent $\delta$ | Number of Roots $P_\delta$ | $\phi(\delta)$ | Number of Periods $S_\delta$ | Number of Primitive Roots $N_\delta$ | Primitive Roots of Congruence $x^\delta \equiv 1$ |
|---|---|---|---|---|---|---|---|
| $2^2.19$ | 18 | 1 | 1 | 1 | 1 | 1 | 1 |
| | | 2 | 4 | 1 | 3 | 3 | 37, 39, 75 |
| | | 3 | 3 | 2 | 1 | 2 | 45, 49 |
| | | 6 | 12 | 2 | 3 | 6 | 7, 11, 27, 31, 65, 69 |
| | | 9 | 9 | 6 | 1 | 6 | 5, 9, 17, 25, 61, 73 |
| | | 18 | 36 | 6 | 3 | 18 | 3, 13, 15, 21, 23, 29, 33, 35, 41, 43, 47, 51, 53, 55, 59, 63, 67, 71 |
| 7.11 | 30 | 1 | 1 | 1 | 1 | 1 | 1 |
| | | 2 | 4 | 1 | 3 | 3 | 34, 43, 76 |
| | | 3 | 3 | 2 | 1 | 2 | 23, 67 |
| | | 5 | 5 | 4 | 1 | 4 | 15, 36, 64, 71 |
| | | 6 | 12 | 2 | 3 | 6 | 10, 12, 32, 45, 54, 65 |
| | | 10 | 20 | 4 | 3 | 12 | 6, 8, 13, 20, 27, 29, 41, 48, 50, 57, 62, 69 |
| | | 15 | 15 | 8 | 1 | 8 | 4, 9, 16, 25, 37, 53, 58, 60 |
| | | 30 | 60 | 8 | 3 | 24 | 2, 3, 5, 17, 18, 19, 24, 26, 30, 31, 38, 39, 40, 46, 47, 51, 52, 59, 61, 68, 72, 73, 74, 75 |
| 2.3.13 | 12 | 1 | 1 | 1 | 1 | 1 | 1 |
| | | 2 | 4 | 1 | 3 | 3 | 25, 53, 77 |
| | | 3 | 3 | 2 | 1 | 2 | 55, 61 |
| | | 4 | 8 | 2 | 2 | 4 | 5, 31, 47, 73 |
| | | 6 | 12 | 2 | 3 | 6 | 17, 23, 29, 35, 43, 49 |
| | | 12 | 24 | 4 | 2 | 8 | 7, 11, 19, 37, 41, 59, 67, 71 |
| 5.17 | 16 | 1 | 1 | 1 | 1 | 1 | 1 |
| | | 2 | 4 | 1 | 3 | 3 | 16, 69, 84 |
| | | 4 | 16 | 2 | 6 | 12 | 4, 13, 18, 21, 33, 38, 47, 52, 64, 67, 72, 81 |
| | | 8 | 32 | 4 | 4 | 16 | 2, 8, 9, 19, 26, 32, 36, 42, 43, 49, 53, 59, 66, 76, 77, 83 |
| | | 16 | 64 | 8 | 4 | 32 | 3, 6, 7, 11, 12, 14, 22, 23, 24, 27, 28, 29, 31, 37, 39, 41, 44, 46, 48, 54, 56, 57, 58, 61, 62, 63, 71, 73, 74, 78, 79, 82 |
| 3.29 | 28 | 1 | 1 | 1 | 1 | 1 | 1 |
| | | 2 | 4 | 1 | 3 | 3 | 28, 59, 86 |
| | | 4 | 8 | 2 | 2 | 4 | 17, 41, 46, 70 |
| | | 7 | 7 | 6 | 1 | 6 | 7, 16, 25, 49, 52, 82 |
| | | 14 | 28 | 6 | 3 | 18 | 4, 5, 13, 20, 22, 23, 34, 35, 38, 53, 62, 64, 65, 67, 71, 74, 80, 83 |
| | | 28 | 56 | 12 | 2 | 24 | 2, 8, 10, 11, 14, 19, 26, 31, 32, 37, 40, 43, 44, 47, 50, 55, 56, 61, 68, 73, 76, 77, 79, 85 |
| $2.3^2.5$ | 12 | 1 | 1 | 1 | 1 | 1 | 1 |
| | | 2 | 4 | 1 | 3 | 3 | 19, 71, 89 |
| | | 3 | 3 | 2 | 1 | 2 | 31, 61 |
| | | 4 | 8 | 2 | 2 | 4 | 17, 37, 53, 73 |
| | | 6 | 12 | 2 | 3 | 6 | 11, 29, 41, 49, 59, 79 |
| | | 12 | 24 | 4 | 2 | 8 | 7, 13, 23, 43, 47, 67, 77, 83 |
| 7.13 | 12 | 1 | 1 | 1 | 1 | 1 | 1 |
| | | 2 | 4 | 1 | 3 | 3 | 27, 64, 90 |
| | | 3 | 9 | 2 | 4 | 8 | 9, 16, 22, 29, 53, 74, 79, 81 |
| | | 4 | 8 | 2 | 2 | 4 | 8, 34, 57, 83 |
| | | 6 | 36 | 2 | 12 | 24 | 3, 4, 10, 12, 17, 23, 25, 30, 36, 38, 40, 43, 48, 51, 55, 61, 62, 66, 68, 69, 75, 82, 87, 88 |
| | | 12 | 72 | 4 | 8 | 32 | 2, 5, 6, 11, 15, 18, 19, 20, 24, 31, 32, 33, 37, 41, 44, 45, 46, 47, 50, 54, 58, 59, 60, 67, 71, 72, 73, 76, 80, 85, 86, 89 |

| Modulus | | Exponent | Number of Roots | | Number of Periods | Number of Primitive Roots | |
|---|---|---|---|---|---|---|---|
| $M$ | $\theta(M)$ | $\delta$ | $P_\delta$ | $\varphi(\delta)$ | $S_\delta$ | $N_\delta$ | Primitive Roots of Congruence $x^\delta \equiv 1$ |
| $2^2.23$ | 22 | 1 | 1 | 1 | 1 | 1 | 1 |
| | | 2 | 4 | 1 | 3 | 3 | 45, 47, 91 |
| | | 11 | 11 | 10 | 1 | 10 | 9, 13, 25, 29, 41, 49, 73, 77, 81, 85 |
| | | 22 | 44 | 10 | 3 | 30 | 3, 5, 7, 11, 15, 17, 19, 21, 27, 31, 33, 35, 37, 39, 43, 51, 53, 55, 57, 59, 61, 63, 65, 67, 71, 75, 79, 83, 87, 89 |
| 3.31 | 30 | 1 | 1 | 1 | 1 | 1 | 1 |
| | | 2 | 4 | 1 | 3 | 3 | 32, 61, 92 |
| | | 3 | 3 | 2 | 1 | 2 | 25, 67 |
| | | 5 | 5 | 4 | 1 | 4 | 4, 16, 64, 70 |
| | | 6 | 12 | 2 | 3 | 6 | 5, 26, 37, 56, 68, 88 |
| | | 10 | 20 | 4 | 3 | 12 | 2, 8, 23, 29, 35, 46, 47, 58, 77, 85, 89, 91 |
| | | 15 | 15 | 8 | 1 | 8 | 7, 10, 19, 28, 40, 49, 76, 82 |
| | | 30 | 60 | 8 | 3 | 24 | 11, 13, 14, 17, 20, 22, 34, 38, 41, 43, 44, 50, 52, 53, 55, 59, 65, 71, 73, 74, 79, 80, 83, 86 |
| 5.19 | 36 | 1 | 1 | 1 | 1 | 1 | 1 |
| | | 2 | 4 | 1 | 3 | 3 | 39, 56, 94 |
| | | 3 | 3 | 2 | 1 | 2 | 11, 26 |
| | | 4 | 8 | 2 | 2 | 4 | 18, 37, 58, 77 |
| | | 6 | 12 | 2 | 3 | 6 | 31, 46, 49, 64, 69, 84 |
| | | 9 | 9 | 6 | 1 | 6 | 6, 16, 36, 61, 66, 81 |
| | | 12 | 24 | 4 | 2 | 8 | 7, 8, 12, 27, 68, 83, 87, 88 |
| | | 18 | 36 | 6 | 3 | 18 | 4, 9, 14, 21, 24, 29, 34, 41, 44, 51, 54, 59, 71, 74, 79, 86, 89, 91 |
| | | 36 | 72 | 12 | 2 | 24 | 2, 3, 13, 17, 22, 23, 28, 32, 33, 42, 43, 47, 48, 52, 53, 62, 63, 67, 72, 73, 78, 82, 92, 93 |
| $3^2.11$ | 30 | 1 | 1 | 1 | 1 | 1 | 1 |
| | | 2 | 4 | 1 | 3 | 3 | 10, 89, 98 |
| | | 3 | 3 | 2 | 1 | 2 | 34, 67 |
| | | 5 | 5 | 4 | 1 | 4 | 37, 64, 82, 91 |
| | | 6 | 12 | 2 | 3 | 6 | 23, 32, 43, 56, 65, 76 |
| | | 10 | 20 | 4 | 3 | 12 | 8, 17, 19, 26, 28, 35, 46, 53, 62, 71, 73, 80 |
| | | 15 | 15 | 8 | 1 | 8 | 4, 16, 25, 31, 49, 58, 70, 97 |
| | | 30 | 60 | 8 | 3 | 24 | 2, 5, 7, 13, 14, 20, 29, 38, 40, 41, 47, 50, 52, 59, 61, 68, 74, 79, 83, 85, 86, 92, 94, 95 |
| $2^2.5^2$ | 20 | 1 | 1 | 1 | 1 | 1 | 1 |
| | | 2 | 4 | 1 | 3 | 3 | 49, 51, 99 |
| | | 4 | 8 | 2 | 2 | 4 | 7, 43, 57, 93 |
| | | 5 | 5 | 4 | 1 | 4 | 21, 41, 61, 81 |
| | | 10 | 20 | 4 | 3 | 12 | 9, 11, 19, 29, 31, 39, 59, 69, 71, 79, 89, 91 |
| | | 20 | 40 | 8 | 2 | 16 | 3, 13, 17, 23, 27, 33, 37, 47, 53, 63, 67, 73, 77, 83, 87, 97 |
| 2.3.17 | 16 | 1 | 1 | 1 | 1 | 1 | 1 |
| | | 2 | 4 | 1 | 3 | 3 | 35, 67, 101 |
| | | 4 | 8 | 2 | 2 | 4 | 13, 47, 55, 89 |
| | | 8 | 16 | 4 | 2 | 8 | 19, 25, 43, 49, 53, 59, 77, 83 |
| | | 16 | 32 | 8 | 2 | 16 | 5, 7, 11, 23, 29, 31, 37, 41, 61, 65, 71, 73, 79, 91, 95, 97 |

| Modulus $M$ | $\theta(M)$ | Exponent $\delta$ | Number of Roots $P_\delta$ | $\phi(\delta)$ | Number of Periods $S_\delta$ | Number of Primitive Roots $N_\delta$ | Primitive Roots of Congruence $x^\delta \equiv 1$ |
|---|---|---|---|---|---|---|---|
| $2^2.3^3$ | 18 | 1 | 1 | 1 | 1 | 1 | 1 |
| | | 2 | 4 | 1 | 3 | 3 | 53, 55, 107 |
| | | 3 | 3 | 2 | 1 | 2 | 37, 73 |
| | | 6 | 12 | 2 | 3 | 6 | 17, 19, 35, 71, 89, 91 |
| | | 9 | 9 | 6 | 1 | 6 | 13, 25, 49, 61, 85, 97 |
| | | 18 | 36 | 6 | 3 | 18 | 5, 7, 11, 23, 29, 31, 41, 43, 47, 59, 65, 67, 77, 79, 83, 95, 101, 107 |
| 2.5.11 | 20 | 1 | 1 | 1 | 1 | 1 | 1 |
| | | 2 | 4 | 1 | 3 | 3 | 21, 89, 109 |
| | | 4 | 8 | 2 | 2 | 4 | 23, 43, 67, 87 |
| | | 5 | 5 | 4 | 1 | 4 | 31, 71, 81, 91 |
| | | 10 | 20 | 4 | 3 | 12 | 9, 19, 29, 39, 41, 49, 51, 59, 61, 69, 79, 101 |
| | | 20 | 40 | 8 | 2 | 16 | 3, 7, 13, 17, 27, 37, 47, 53, 57, 63, 73, 83, 93, 97, 103, 107 |
| 3.37 | 36 | 1 | 1 | 1 | 1 | 1 | 1 |
| | | 2 | 4 | 1 | 3 | 3 | 38, 73, 110 |
| | | 3 | 3 | 2 | 1 | 2 | 10, 100 |
| | | 4 | 8 | 2 | 2 | 4 | 31, 43, 68, 80 |
| | | 6 | 12 | 2 | 3 | 6 | 11, 26, 47, 64, 85, 101 |
| | | 9 | 9 | 6 | 1 | 6 | 7, 16, 34, 46, 49, 70 |
| | | 12 | 24 | 4 | 2 | 8 | 8, 14, 23, 29, 82, 88, 97, 103 |
| | | 18 | 36 | 6 | 3 | 18 | 4, 25, 28, 40, 41, 44, 53, 58, 62, 65, 67, 71, 77, 83, 86, 95, 104, 107 |
| | | 36 | 72 | 12 | 2 | 24 | 2, 5, 13, 17, 19, 20, 22, 32, 35, 50, 52, 55, 56, 59, 61, 76, 79, 89, 91, 92, 94, 98, 106, 109 |
| 2.3.19 | 18 | 1 | 1 | 1 | 1 | 1 | 1 |
| | | 2 | 4 | 1 | 3 | 3 | 37, 77, 113 |
| | | 3 | 3 | 2 | 1 | 2 | 7, 49 |
| | | 6 | 12 | 2 | 3 | 6 | 11, 31, 65, 83, 103, 107 |
| | | 9 | 9 | 6 | 1 | 6 | 25, 43, 55, 61, 73, 85 |
| | | 18 | 36 | 6 | 3 | 18 | 5, 13, 17, 23, 29, 35, 41, 47, 53, 59, 67, 71, 79, 89, 91, 97, 101, 109 |
| 5.23 | 44 | 1 | 1 | 1 | 1 | 1 | 1 |
| | | 2 | 4 | 1 | 3 | 3 | 24, 91, 114 |
| | | 4 | 8 | 2 | 2 | 4 | 22, 47, 68, 93 |
| | | 11 | 11 | 10 | 1 | 10 | 6, 16, 26, 31, 36, 41, 71, 81, 96, 101 |
| | | 22 | 44 | 10 | 3 | 30 | 4, 9, 11, 14, 19, 21, 29, 34, 39, 44, 49, 51, 54, 56, 59, 61, 64, 66, 74, 76, 79, 84, 86, 89, 94, 99, 104, 106, 109, 111 |
| | | 44 | 88 | 20 | 2 | 40 | 2, 3, 7, 8, 12, 13, 17, 18, 27, 28, 32, 33, 37, 38, 42, 43, 48, 52, 53, 57, 58, 62, 63, 67, 72, 73, 77, 78, 82, 83, 87, 88, 97, 98, 102, 103, 107, 108, 112, 113 |
| $2^2.29$ | 28 | 1 | 1 | 1 | 1 | 1 | 1 |
| | | 2 | 4 | 1 | 3 | 3 | 57, 59, 115 |
| | | 4 | 8 | 2 | 2 | 4 | 17, 41, 75, 99 |
| | | 7 | 7 | 6 | 1 | 6 | 25, 45, 49, 53, 65, 81 |
| | | 14 | 28 | 6 | 3 | 18 | 5, 7, 9, 13, 23, 33, 35, 51, 63, 67, 71, 83, 91, 93, 103, 107, 109, 111 |
| | | 28 | 56 | 12 | 2 | 24 | 3, 11, 15, 19, 21, 27, 31, 37, 39, 43, 47, 55, 61, 69, 73, 77, 79, 85, 89, 95, 97, 101, 105, 113 |

| Modulus | | Exponent | Number of Roots | | Number of Periods | Number of Primitive Roots | |
|---|---|---|---|---|---|---|---|
| $M$ | $\theta(M)$ | $\delta$ | $P_\delta$ | $\phi(\delta)$ | $S_\delta$ | $N_\delta$ | Primitive Roots of Congruence $x^\delta \equiv 1$ |
| $3^2.13$ | 12 | 1 | 1 | 1 | 1 | 1 | 1 |
| | | 2 | 4 | 1 | 3 | 3 | 53, 64, 116 |
| | | 3 | 9 | 2 | 4 | 8 | 16, 22, 40, 55, 61, 79, 94, 100 |
| | | 4 | 8 | 2 | 2 | 4 | 8, 44, 73, 109 |
| | | 6 | 36 | 2 | 12 | 24 | 4, 10, 14, 17, 23, 25, 29, 35, 38, 43, 49, 56, 62, 68, 74, 77, 82, 88, 92, 95, 101, 103, 107, 113 |
| | | 12 | 72 | 4 | 8 | 32 | 2, 5, 7, 11, 19, 20, 28, 31, 32, 34, 37, 41, 46, 47, 50, 58, 59, 67, 70, 71, 76, 80, 83, 85, 86, 89, 97, 98, 106, 110, 112, 115 |
| 7.17 | 48 | 1 | 1 | 1 | 1 | 1 | 1 |
| | | 2 | 4 | 1 | 3 | 3 | 50, 69, 118 |
| | | 3 | 3 | 2 | 1 | 2 | 18, 86 |
| | | 4 | 8 | 2 | 2 | 4 | 13, 55, 64, 106 |
| | | 6 | 12 | 2 | 3 | 6 | 16, 33, 52, 67, 101, 103 |
| | | 8 | 16 | 4 | 2 | 8 | 8, 15, 36, 43, 76, 83, 104, 111 |
| | | 12 | 24 | 4 | 2 | 8 | 4, 30, 38, 47, 72, 81, 89, 115 |
| | | 16 | 32 | 8 | 2 | 16 | 6, 20, 22, 27, 29, 41, 48, 57, 62, 71, 78, 90, 92, 97, 99, 113 |
| | | 24 | 48 | 8 | 2 | 16 | 2, 9, 19, 25, 26, 32, 53, 59, 60, 66, 87, 93, 94, 100, 110, 117 |
| | | 48 | 96 | 16 | 2 | 32 | 3, 5, 10, 11, 12, 23, 24, 31, 37, 39, 40, 44, 45, 46, 54, 58, 61, 65, 73, 74, 75, 79, 80, 82, 88, 95, 96, 107, 108, 109, 114, 116 |

## THE THIRD CATEGORY

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| $2^3.3$ | 2 | 1 | 1 | 1 | 1 | 1 | 1 |
| | | 2 | 8 | 1 | 7 | 7 | 2, 5, 7, 11, 13, 17, 19, 23 |
| $2^3.5$ | 4 | 1 | 1 | 1 | 1 | 1 | 1 |
| | | 2 | 8 | 1 | 7 | 7 | 9, 11, 19, 21, 29, 31, 39 |
| | | 4 | 16 | 2 | 4 | 8 | 3, 7, 13, 17, 23, 27, 33, 37 |
| $2^4.3$ | 4 | 1 | 1 | 1 | 1 | 1 | 1 |
| | | 2 | 8 | 1 | 7 | 7 | 7, 17, 23, 25, 31, 41, 47 |
| | | 4 | 16 | 2 | 4 | 8 | 5, 11, 13, 19, 29, 35, 37, 43 |
| $2^3.7$ | 6 | 1 | 1 | 1 | 1 | 1 | 1 |
| | | 2 | 8 | 1 | 7 | 7 | 13, 15, 27, 29, 41, 43, 55 |
| | | 3 | 3 | 2 | 1 | 2 | 9, 25 |
| | | 6 | 24 | 2 | 7 | 14 | 3, 5, 11, 17, 19, 23, 31, 33, 37, 39, 45, 47, 51, 53 |
| $2^2.3.5$ | 4 | 1 | 1 | 1 | 1 | 1 | 1 |
| | | 2 | 8 | 1 | 7 | 7 | 11, 19, 29, 31, 41, 49, 59 |
| | | 4 | 16 | 2 | 4 | 8 | 7, 13, 17, 23, 37, 43, 47, 53 |

| Modulus $M$ | $\theta(M)$ | Exponent $\delta$ | Number of Roots $P_\delta$ | $\phi(\delta)$ | Number of Periods $S_\delta$ | Number of Primitive Roots $N_\delta$ | Primitive Roots of Congruence $x^\delta \equiv 1$ |
|---|---|---|---|---|---|---|---|
| $2^3.3^2$ | 6 | 1 | 1 | 1 | 1 | 1 | 1 |
| | | 2 | 8 | 1 | 7 | 7 | 17, 19, 35, 37, 53, 55, 71 |
| | | 3 | 3 | 2 | 1 | 2 | 25, 49 |
| | | 6 | 24 | 2 | 7 | 14 | 5, 7, 11, 13, 23, 29, 31, 41, 43, 47, 59, 61, 65, 67 |
| $2^4.5$ | 4 | 1 | 1 | 1 | 1 | 1 | 1 |
| | | 2 | 8 | 1 | 7 | 7 | 9, 31, 39, 41, 49, 71, 79 |
| | | 4 | 32 | 2 | 12 | 24 | 3, 7, 11, 13, 17, 19, 21, 23, 27, 29, 33, 37, 43, 47, 51, 53, 57, 59, 61, 63, 67, 69, 73, 77 |
| $2^2.3.7$ | 6 | 1 | 1 | 1 | 1 | 1 | 1 |
| | | 2 | 8 | 1 | 7 | 7 | 13, 29, 41, 43, 55, 71, 83 |
| | | 3 | 3 | 2 | 1 | 2 | 25, 37 |
| | | 6 | 24 | 2 | 7 | 14 | 5, 11, 17, 19, 23, 31, 47, 53, 59, 61, 65, 67, 73, 79 |
| $2^3.11$ | 10 | 1 | 1 | 1 | 1 | 1 | 1 |
| | | 2 | 8 | 1 | 7 | 7 | 21, 23, 43, 45, 65, 67, 87 |
| | | 5 | 5 | 4 | 1 | 4 | 9, 25, 49, 81 |
| | | 10 | 40 | 4 | 7 | 28 | 3, 5, 7, 13, 15, 17, 19, 27, 29, 31, 35, 37, 39, 41, 47, 51, 53, 57, 59, 61, 63, 69, 71, 73, 75, 79, 83, 85 |
| $2^5.3$ | 8 | 1 | 1 | 1 | 1 | 1 | 1 |
| | | 2 | 8 | 1 | 7 | 7 | 17, 31, 47, 49, 65, 79, 95 |
| | | 4 | 16 | 2 | 4 | 8 | 7, 23, 25, 41, 55, 71, 73, 89 |
| | | 8 | 32 | 4 | 4 | 16 | 5, 11, 13, 19, 29, 35, 37, 43, 53, 59, 61, 67, 77, 83, 85, 91 |
| $2^3.13$ | 12 | 1 | 1 | 1 | 1 | 1 | 1 |
| | | 2 | 8 | 1 | 7 | 7 | 25, 27, 51, 53, 77, 79, 103 |
| | | 3 | 3 | 2 | 1 | 2 | 9, 81 |
| | | 4 | 16 | 2 | 4 | 8 | 5, 21, 31, 47, 57, 73, 83, 99 |
| | | 6 | 24 | 2 | 7 | 14 | 3, 17, 23, 29, 35, 43, 49, 55, 61, 69, 75, 87, 95, 101 |
| | | 12 | 48 | 4 | 4 | 16 | 7, 11, 15, 19, 33, 37, 41, 45, 59, 63, 67, 71, 85, 89, 93, 97 |
| 3.5.7 | 12 | 1 | 1 | 1 | 1 | 1 | 1 |
| | | 2 | 8 | 1 | 7 | 7 | 29, 34, 41, 64, 71, 76, 104 |
| | | 3 | 3 | 2 | 1 | 2 | 16, 46 |
| | | 4 | 16 | 2 | 4 | 8 | 8, 13, 22, 43, 62, 83, 92, 97 |
| | | 6 | 24 | 2 | 7 | 14 | 4, 11, 19, 26, 31, 44, 59, 61, 74, 79, 86, 89, 94, 101 |
| | | 12 | 48 | 4 | 4 | 16 | 2, 17, 23, 32, 37, 38, 47, 52, 53, 58, 67, 68, 73, 82, 88, 103 |
| $2^4.7$ | 12 | 1 | 1 | 1 | 1 | 1 | 1 |
| | | 2 | 8 | 1 | 7 | 7 | 15, 41, 55, 57, 71, 97, 111 |
| | | 3 | 3 | 2 | 1 | 2 | 65, 81 |
| | | 4 | 16 | 2 | 4 | 8 | 13, 27, 29, 43, 69, 83, 85, 99 |
| | | 6 | 24 | 2 | 7 | 14 | 9, 17, 23, 25, 31, 33, 39, 47, 73, 79, 87, 89, 95, 103 |
| | | 12 | 48 | 4 | 4 | 16 | 3, 5, 11, 19, 37, 45, 51, 53, 59, 61, 67, 75, 93, 101, 107, 109 |

# XXII. *On the Stability of the Steady Motion of Viscous Liquid contained between two Rotating Coaxal Circular Cylinders.*

By K. Tamaki, Assistant Professor of Mathematics, Kyoto Imperial University and W. J. Harrison, M.A., Fellow of Clare College, Cambridge.

[*Received* 27 July, 1920.]

*Introductory Statement.*

(*a*) The first part of this paper is occupied with a discussion of the stability of the motion cited in the title. Mallock's experiments* in this connection are well known, and Lord Rayleigh† has given an interesting account of Lord Kelvin's observations on them. In Mallock's viscometer the outer cylinder was made to rotate, as it was found that the motion of the liquid was always turbulent when the inner cylinder was rotated.

A theoretical investigation of the stability on the lines of the investigations carried out by Reynolds and Orr shows that there is no difference in the relative stabilities of the two cases, provided that in one case the inner cylinder has an angular velocity equal to that of the outer in the other case. The observed difference in stability must therefore be accounted for by making the hypothesis that the same type of disturbance is not likely to be set up in both cases, but that the disturbances which arise in the one case are more likely to cause instability than the disturbances which arise in the other. Some investigations on this point are given in Part II. It is clear that this discriminative action is a matter for hypothesis, or for experimental verification; it cannot be inferred from the hydrodynamical equations.

Criteria are obtained analogous to those given by Reynolds and Orr for other types of motion.

(*b*) In the second part of this paper, in addition to the investigations mentioned above, a note is given on a criterion suggested by Lamb.

An investigation is also made into the effect of proceeding to a higher approximation on the usual criterion for stability. It is thought that an explanation may thus be found of the discrepancy between the conclusion arrived at by Reynolds that a certain degree of viscosity is necessary for stability, and the conclusion of Lord Rayleigh that certain steady motions of a non-viscous liquid are stable.

* *Proc. Roy. Soc.*, vol. XLV. p. 126, 1888.

† *Phil. Mag.* (6), vol. XXVIII. p. 610, 1914.

## PART I.

### By K. TAMAKI.

Since Osborne Reynolds found that the steady motion of viscous fluid passing through a circular pipe becomes unstable at a certain critical mean velocity of flow, the question of stability of the steady motion of fluid has been attacked theoretically by several eminent authors. Professor Orr discussed this problem thoroughly in his paper, "The Stability or Instability of the Steady Motion of a Perfect Liquid and of a Viscous Liquid*." He obtained the minimum critical velocities in the cases of flow between two parallel planes and flow through a circular pipe by substituting some differential equations for Reynolds' discriminating equation in integral form determining the critical velocity. So far as the writer is aware the case in which one of two concentric cylinders is rotating steadily while the other is fixed has not yet been investigated. The aim of this paper is to investigate this case, applying Professor Orr's method.

If we confine ourselves to a consideration of the two-dimensional motions of a viscous fluid we have the following equations of motion, namely

$$\left.\begin{aligned}\frac{\partial u}{\partial t}+u\frac{\partial u}{\partial x}+v\frac{\partial u}{\partial y}&=\frac{1}{\rho}\left(\frac{\partial p_{xx}}{\partial x}+\frac{\partial p_{xy}}{\partial y}\right)\\ \frac{\partial v}{\partial t}+u\frac{\partial v}{\partial x}+v\frac{\partial v}{\partial y}&=\frac{1}{\rho}\left(\frac{\partial p_{yx}}{\partial x}+\frac{\partial p_{yy}}{\partial y}\right)\end{aligned}\right\} \qquad \text{..............................(1)},$$

where $\rho$ is the density of the fluid, $u$, $v$ the components of velocity and $p_{xx}$, $p_{xy}$, $p_{yy}$ the components of stress.

Let $U$, $V$ be the components of velocity and $P_{xx}$, $P_{xy}$, $P_{yy}$ the components of stress of the fluid when its motion is steady, then by (1) we have

$$\left.\begin{aligned}U\frac{\partial U}{\partial x}+V\frac{\partial U}{\partial y}&=\frac{1}{\rho}\left(\frac{\partial P_{xx}}{\partial x}+\frac{\partial P_{xy}}{\partial y}\right)\\ U\frac{\partial V}{\partial x}+V\frac{\partial V}{\partial y}&=\frac{1}{\rho}\left(\frac{\partial P_{yx}}{\partial x}+\frac{\partial P_{yy}}{\partial y}\right)\end{aligned}\right\} \qquad \text{..............................(2)}.$$

Now let $u$, $v$ be the components of velocity of the turbulent motion and $p_{xx}$, $p_{xy}$, $p_{yy}$ the components of stress due to it so that $U+u$, $V+v$, and $P_{xx}+p_{xx}$, $P_{xy}+p_{xy}$, $P_{yy}+p_{yy}$ are the components of velocity and of stress respectively when the motion of the liquid is disturbed from its steady state. Then substituting these quantities in (1), we get

$$\frac{\partial u}{\partial t}+(U+u)\frac{\partial}{\partial x}(U+u)+(V+v)\frac{\partial}{\partial y}(U+u)=\frac{1}{\rho}\left[\frac{\partial}{\partial x}(P_{xx}+p_{xx})+\frac{\partial}{\partial y}(P_{xy}+p_{xy})\right],$$

$$\frac{\partial v}{\partial t}+(U+u)\frac{\partial}{\partial x}(V+v)+(V+v)\frac{\partial}{\partial y}(V+v)=\frac{1}{\rho}\left[\frac{\partial}{\partial x}(P_{xy}+p_{xy})+\frac{\partial}{\partial y}(P_{yy}+p_{yy})\right].$$

Subtracting equations (2) from the corresponding equations above written and neglecting the squares and products of $u$, $v$ we get

$$\frac{\partial u}{\partial t}+u\frac{\partial U}{\partial x}+v\frac{\partial U}{\partial y}+U\frac{\partial u}{\partial x}+V\frac{\partial u}{\partial y}=\frac{1}{\rho}\left(\frac{\partial p_{xx}}{\partial x}+\frac{\partial p_{xy}}{\partial y}\right) \qquad \text{.....................(3)},$$

$$\frac{\partial v}{\partial t}+u\frac{\partial V}{\partial x}+v\frac{\partial V}{\partial y}+U\frac{\partial v}{\partial x}+V\frac{\partial v}{\partial y}=\frac{1}{\rho}\left(\frac{\partial p_{yx}}{\partial x}+\frac{\partial p_{yy}}{\partial y}\right) \qquad \text{.....................(4)}.$$

* *Proc. Royal Irish Acad.*, vol. XXVII. Sect. A, p. 9, 1907.

Multiplying (3) and (4) by $\rho u$, $\rho v$ respectively, adding and integrating throughout the volume of any portion of the fluid we get

$$\frac{\partial}{\partial t}\frac{1}{2}\int \rho\,(u^2+v^2)\,d\tau = -\int \rho u\left(u\frac{\partial U}{\partial x}+v\frac{\partial U}{\partial y}\right)d\tau - \int \rho v\left(u\frac{\partial V}{\partial x}+v\frac{\partial V}{\partial y}\right)d\tau$$
$$-\int \rho u\left(U\frac{\partial u}{\partial x}+V\frac{\partial u}{\partial y}\right)d\tau - \int \rho v\left(U\frac{\partial v}{\partial x}+V\frac{\partial v}{\partial y}\right)d\tau$$
$$+\int\left[u\left(\frac{\partial p_{xx}}{\partial x}+\frac{\partial p_{xy}}{\partial y}\right)+v\left(\frac{\partial p_{yx}}{\partial x}+\frac{\partial p_{yy}}{\partial y}\right)\right]d\tau,$$

where $d\tau$ denotes an elementary volume.

Integrating by parts the integrals on the right-hand side of the above equation we get

$$\frac{\partial}{\partial t}\frac{1}{2}\int \rho\,(u^2+v^2)\,d\tau = -\frac{1}{2}\int \rho\,(u^2+v^2)\,(Ul+Vm)\,dS$$
$$+\int[u\,(p_{xx}l+p_{xy}m)+v\,(p_{yx}l+p_{yy}m)]\,dS$$
$$-\int\rho\left[u\left(u\frac{\partial}{\partial x}+v\frac{\partial}{\partial y}\right)U+v\left(u\frac{\partial}{\partial x}+v\frac{\partial}{\partial y}\right)V\right]d\tau$$
$$+\frac{1}{2}\int\rho\,(u^2+v^2)\left(\frac{\partial U}{\partial x}+\frac{\partial V}{\partial y}\right)d\tau$$
$$-\int\left[p_{xx}\frac{\partial u}{\partial x}+p_{xy}\frac{\partial v}{\partial y}+p_{yy}\left(\frac{\partial u}{\partial x}+\frac{\partial v}{\partial y}\right)\right]d\tau,$$

where $l$, $m$ are the direction cosines of the normal drawn outwards to the surface element $dS$. The first integral on the right-hand side of this equation represents the rate at which the kinetic energy of the turbulent motion is convected into the volume considered and the second integral expresses the time rate at which the stress due to the turbulent motion does work on the fluid in this volume. In some cases we can choose the boundary surfaces so as to make the joint effect of these terms cancel. For example, let us suppose that the fluid is flowing between two parallel planes $y=\pm b$ and the turbulent motion has a wave-length $a$ with respect to $x$. Then by taking $y=\pm b$, $x=x$, $x=x+a$ as the boundary surfaces we can satisfy the condition of the cancelling of these terms. Therefore let us suppose that we have chosen a closed surface so as to cancel this joint effect. Then we have

$$\frac{\partial}{\partial t}\frac{1}{2}\int \rho\,(u^2+v^2)\,d\tau = -\int\rho\left[u\left(u\frac{\partial}{\partial x}+v\frac{\partial}{\partial y}\right)U+v\left(u\frac{\partial}{\partial x}+v\frac{\partial}{\partial y}\right)V\right]d\tau$$
$$+\frac{1}{2}\int\rho\,(u^2+v^2)\left(\frac{\partial U}{\partial x}+\frac{\partial V}{\partial y}\right)d\tau$$
$$-\int\left[p_{xx}\frac{\partial u}{\partial x}+p_{xy}\frac{\partial v}{\partial y}+p_{yy}\left(\frac{\partial u}{\partial x}+\frac{\partial v}{\partial y}\right)\right]d\tau \quad \ldots\ldots\ldots\ldots(5).$$

Since we have

$$\left.\begin{aligned} p_{xx} &= -p-\frac{2}{3}\operatorname{div} q+2\frac{\partial u}{\partial x}\\ p_{yy} &= -p-\frac{2}{3}\operatorname{div} q+2\frac{\partial v}{\partial y}\\ p_{xy} &= \mu\left(\frac{\partial u}{\partial x}+\frac{\partial v}{\partial y}\right)\end{aligned}\right\} \quad \ldots\ldots\ldots\ldots(6),$$

where $\mu$ is the coefficient of viscosity and $q$ the velocity vector of the turbulent motion, (5) may be written

$$\begin{aligned}\frac{\partial}{\partial t}\frac{1}{2}\int \rho q^2\,d\tau = &-\int\rho\left[u\left(u\frac{\partial U}{\partial x}+v\frac{\partial U}{\partial y}\right)+v\left(u\frac{\partial V}{\partial x}+v\frac{\partial V}{\partial y}\right)\right]d\tau\\ &+\frac{1}{2}\int\rho q^2\operatorname{div} Q\,d\tau\\ &-\mu\int\left[2\left(\frac{\partial u}{\partial x}\right)^2+2\left(\frac{\partial v}{\partial y}\right)^2+\left(\frac{\partial u}{\partial x}+\frac{\partial v}{\partial y}\right)^2\right]d\tau\\ &+\int p'\operatorname{div} q\,d\tau \quad \text{.........................................}(7),\end{aligned}$$

where $Q$ denotes the velocity of the steady motion and

$$p'=p+\frac{2}{3}\operatorname{div} q \quad \text{.........................................}(8).$$

The expression (7) gives the time rate of change of energy of the turbulent motion and we see that whether the disturbance from the steady motion increases or decreases depends wholly on the sign of the whole expression. Hence if for a given steady motion we could find the lowest limit of $\mu$ for which it is possible to choose $q$ so as to make the expression (7) zero we could find the critical velocity for a given value of $\mu$.

Now putting the right-hand side of equation (7) equal to zero and integrating by parts the integrals except the second, we get

$$\begin{aligned}&-\int\frac{1}{2}\rho\left[u\left\{\left(u\frac{\partial U}{\partial x}+v\frac{\partial U}{\partial y}\right)+\left(u\frac{\partial U}{\partial x}+v\frac{\partial V}{\partial y}\right)\right\}+v\left\{\left(u\frac{\partial V}{\partial x}+v\frac{\partial V}{\partial y}\right)+\left(u\frac{\partial U}{\partial x}+v\frac{\partial V}{\partial y}\right)\right\}\right]d\tau\\ &+\int\frac{1}{2}\rho\left(u^2+v^2\right)\operatorname{div} Q\,d\tau\\ &+\int\left[u\left(\nabla^2 u+\frac{\partial}{\partial x}\operatorname{div} q\right)+v\left(\nabla^2 v+\frac{\partial}{\partial y}\operatorname{div} q\right)\right]d\tau\\ &-\int\left(u\frac{\partial p'}{\partial x}+v\frac{\partial p'}{\partial y}\right)d\tau=0.\end{aligned}$$

where
$$\nabla^2=\frac{\partial^2}{\partial x^2}+\frac{\partial^2}{\partial y^2}.$$

Hence we must have

$$-\frac{1}{2}\rho\left[\left(u\frac{\partial U}{\partial x}+v\frac{\partial U}{\partial y}\right)+\left(u\frac{\partial U}{\partial x}+v\frac{\partial V}{\partial y}\right)-u\operatorname{div} Q\right]+\mu\left(\nabla^2 u+\frac{\partial}{\partial x}\operatorname{div} q\right)=\frac{\partial p'}{\partial x},$$

$$-\frac{1}{2}\rho\left[\left(u\frac{\partial V}{\partial x}+v\frac{\partial V}{\partial y}\right)+\left(u\frac{\partial U}{\partial x}+v\frac{\partial V}{\partial y}\right)-v\operatorname{div} Q\right]+\mu\left(\nabla^2 v+\frac{\partial}{\partial y}\operatorname{div} q\right)=\frac{\partial p'}{\partial y}.$$

If the fluid is incompressible these equations simplify into

$$\left.\begin{aligned}2\mu\nabla^2 u-\rho\left[\left(u\frac{\partial U}{\partial x}+v\frac{\partial U}{\partial y}\right)+\left(u\frac{\partial U}{\partial x}+v\frac{\partial V}{\partial y}\right)\right]=2\frac{\partial p}{\partial x}\\ 2\mu\nabla^2 v-\rho\left[\left(u\frac{\partial V}{\partial x}+v\frac{\partial V}{\partial y}\right)+\left(u\frac{\partial U}{\partial x}+v\frac{\partial V}{\partial y}\right)\right]=2\frac{\partial p}{\partial y}\end{aligned}\right\} \quad \text{..................}(9).$$

These are Prof. Orr's equations extended for the case in which the steady motion of the fluid has $y$-component as well as $x$-component of velocity.

It is well known that when one of the concentric cylinders is rotating with a constant velocity the stream-lines are concentric circles and the velocity is given by

$$V = \frac{A}{r} + Br \quad \text{.........................................(10)},$$

where $A$ and $B$ are constants determined by boundary conditions. Hence transforming from the rectangular coordinates $x, y$ to the polar coordinates $r, \theta$ and expressing the radial and tangential components of the turbulent motion by $u$ and $v$ respectively, equations (9) may be written

$$\left.\begin{aligned} 2\mu\nabla^2 u - \frac{2\mu}{r}\left(u + 2\frac{\partial v}{\partial\theta}\right) - \rho v\left(\frac{\partial V}{\partial r} - \frac{V}{r}\right) = 2\frac{\partial p}{\partial r} \\ 2\mu\nabla^2 v - \frac{2\mu}{r}\left(v + 2\frac{\partial u}{\partial\theta}\right) - \rho u\left(\frac{\partial V}{\partial r} - \frac{V}{r}\right) = 2\frac{1}{r}\frac{\partial p}{\partial\theta} \end{aligned}\right\} \quad \text{....................(11)}.$$

Eliminating $p$ between these equations we get

$$2\mu\nabla^2\left(\frac{\partial v}{\partial r} + \frac{v}{r} - \frac{1}{r}\frac{\partial u}{\partial\theta}\right) - \rho\left[u\frac{\partial F}{\partial r} + F\left(\frac{u}{r} + \frac{\partial u}{\partial r} - \frac{1}{r}\frac{\partial v}{\partial\theta}\right)\right] = 0 \quad \text{..............(12)},$$

where we put

$$F = \frac{\partial V}{\partial r} - \frac{V}{r} \text{.........................................(13)}.$$

Making use of the equation of continuity

$$\frac{\partial u}{\partial r} + \frac{u}{r} + \frac{1}{r}\frac{\partial v}{\partial\theta} = 0 \quad \text{.........................................(14)},$$

and substituting the value of $V$ given by (10), this equation becomes

$$\mu\nabla^2\left(\frac{\partial v}{\partial r} + \frac{v}{r} - \frac{1}{r}\frac{\partial u}{\partial\theta}\right) - \frac{2\rho A}{r^3}\left(u + \frac{\partial v}{\partial\theta}\right) = 0 \text{..........................(15)}.$$

From (14) we see that we may use the stream function $\psi$ defined by

$$u = -\frac{1}{r}\frac{\partial\psi}{\partial\theta}, \qquad v = \frac{\partial\psi}{\partial r} \quad \text{....................................(16)}.$$

From the nature of the problem it is natural to assume that $\psi$ varies as $e^{i\lambda\theta}$, where $\lambda$ denotes an unknown integer. Then (15) may be written

$$\nabla^4\psi + \frac{k}{r^3}\left(\frac{1}{r} - \frac{\partial}{\partial r}\right)\psi = 0 \quad \text{....................................(17)},$$

where

$$k = i\frac{2\lambda\rho A}{\mu} \quad \text{..........................................(18)}.$$

Expanding $\nabla^4\psi$, (17) becomes

$$\frac{d^4\psi}{dr^4} + \frac{2}{r}\frac{d^3\psi}{dr^3} - \frac{1+2\lambda^2}{r^2}\frac{d^2\psi}{dr^2} + \frac{1-k+2\lambda^2}{r^3}\frac{d\psi}{dr} + \frac{k-4\lambda^2+\lambda^4}{r^4}\psi = 0 \text{ ...........(19)}.$$

This is an homogeneous linear differential equation of the fourth order. Therefore assuming

$$\psi = r^{m+1} \text{ ............................................(20)},$$

we get the biquadratic equation

$$m^4 - 2(1+\lambda^2)m^2 - km + (1-\lambda^2)^2 = 0 \text{ ..........................(21)},$$

to determine the values of $m$.

Since the coefficient of $m^3$ is zero we can assume as the roots of this equation

$$m_1 = p + s, \ m_2 = p - s, \ m_3 = -p + \sigma, \ m_4 = -p - \sigma \text{....................(22)}.$$

Then the equation (21) gives the following relations

$$\left.\begin{aligned} 2p^2 + s^2 + \sigma^2 &= 2(1+\lambda^2) \\ 2p(s^2 - \sigma^2) &= k \\ (p^2 - s^2)(p^2 - \sigma^2) &= (1-\lambda^2)^2 \end{aligned}\right\} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots(23).$$

Since the disturbed motion must have the same velocity on the surfaces of the cylinders $r = a$ and $r = b$ as that of the steady motion, assuming that the angular velocities of the cylinders are maintained constant, the velocity of the turbulent motion must vanish on these surfaces. Hence the boundary conditions are $\psi = 0$, $\frac{\partial \psi}{\partial r} = 0$ when $r = a$ and $r = b$. Putting

$$\psi = a_1 r^{m_1+1} + a_2 r^{m_2+1} + a_3 r^{m_3+1} + a_4 r^{m_4+1} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots(24),$$

where $a_1$, $a_2$, $a_3$, $a_4$ are constants as yet undetermined, the boundary conditions become

$$a_1 a^{m_1} + a_2 a^{m_2} + a_3 a^{m_3} + a_4 a^{m_4} = 0,$$

$$a_1 b^{m_1} + a_2 b^{m_2} + a_3 b^{m_3} + a_4 b^{m_4} = 0,$$

$$a_1 m_1 a^{m_1} + a_2 m_2 a^{m_2} + a_3 m_3 a^{m_3} + a_4 m_4 a^{m_4} = 0,$$

$$a_1 m_1 b^{m_1} + a_2 m_2 b^{m_2} + a_3 m_3 b^{m_3} + a_4 m_4 b^{m_4} = 0.$$

Eliminating $a_1$, $a_2$, $a_3$, $a_4$ and substituting the values of $m_1$, $m_2$, $m_3$, $m_4$ we get

$$\begin{aligned} &[4p^2 - (\sigma - s)^2][a^{\sigma+s} b^{-(\sigma+s)} + a^{-(\sigma+s)} b^{\sigma+s}] \\ &- [4p^2 - (\sigma + s)^2][a^{\sigma-s} b^{-(\sigma-s)} + a^{-(\sigma-s)} b^{\sigma-s}] \\ &- 4\sigma s (a^{2p} b^{-2p} + a^{-2p} b^{2p}) = 0 \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots(25). \end{aligned}$$

This equation is satisfied by putting*

$$2p = \sigma + s \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots(26).$$

In order that this relation (26) may be consistent with the first two equations of (23) without destroying the relation

$$m_1 + m_2 + m_3 + m_4 = 0,$$

we must assume

$$\sigma = i\alpha + \beta, \quad s = -i\alpha + \gamma \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots(27).$$

Substituting these values of $\sigma$ and $s$ into the first equation of (23) we have

$$2(p^2 - \alpha^2) + \beta^2 + \gamma^2 + 2i\alpha(\beta - \gamma) = 2(1 + \lambda^2).$$

Since the right-hand side of this equation is real we must have

$$\beta = \gamma$$

as far as we consider $p$, $\alpha$, $\beta$, $\gamma$ as real quantities. Then (26) and (27) together with this relation give us

$$\beta = \gamma = p,$$

and consequently we have

$$\sigma = p + i\alpha, \quad s = p - i\alpha \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots(28).$$

* Equation (25) can be satisfied by assuming $2p = \sigma + s$, $2p = \sigma - s$, $2p = -(\sigma + s)$, $2p = -(\sigma - s)$. But we can easily see that we arrive at the same result by making any one of these alternative assumptions.

Substituting these values of $\sigma$ and $s$ into (23) we get

$$-8i\alpha p^2 = k \qquad \text{(29)},$$

$$2p^2 - \alpha^2 = 1 + \lambda^2 \qquad \text{(30)},$$

$$\alpha^2(\alpha^2 + 4p^2) = (1 - \lambda^2)^2 \qquad \text{(31)}.$$

Solving for $\alpha^2$ and $p^2$ we obtain

$$\alpha^2 = \frac{1}{3}[-(1 + \lambda^2) \pm 2\sqrt{1 - \lambda^2 + \lambda^4}] \qquad \text{(32)},$$

$$p^2 = \frac{1}{3}[1 + \lambda^2 \pm \sqrt{1 - \lambda^2 + \lambda^4}] \qquad \text{(33)}.$$

If we put $\lambda = 1$ in (32) we have

$$\alpha^2 = \frac{1}{3}(-2 \pm 2),$$

therefore we see that we must take the upper sign in (32) and (33).

Substituting the values of $k$, $\alpha$ and $p$ in (29) we obtain

$$\mu = \pm \frac{3\sqrt{3}}{4} \frac{\lambda \rho A}{[1 + \lambda^2 + \sqrt{1 - \lambda^2 + \lambda^4}][-(1 + \lambda^2) + 2\sqrt{1 - \lambda^2 + \lambda^4}]^{\frac{1}{2}}} \qquad \text{(34)},$$

the positive or negative sign must be taken according as $A$ is positive or negative. As $\lambda$ increases from zero to $+1$, $\mu$ increases from zero to infinity and then it gradually decreases to zero as $\lambda$ increases. For a disturbance which is independent of $\theta$ the motion is accordingly always stable. For one particular disturbance depending on $\cos(\theta + \epsilon)$ only, as regards $\theta$, the motion is unstable for all values of $\mu$*. Taking $\lambda = 2$, we get

$$\frac{\rho A}{\mu} = \pm 4{\cdot}924 \qquad \text{(35)}.$$

This corresponds to a much lower critical value of the velocity than the critical values obtained by other authors in the cases of flow between parallel planes and through a circular pipe. As $\lambda$ is increased the critical value of $\frac{\rho A}{\mu}$ is increased. Thus there is an inherent instability for any value of $\mu$ if a particular type of disturbance is set up by itself. The difficulty which arises lies in the fact that this instability is only in evidence when the inner cylinder is rotated.

When one of the cylinders is rotating with a constant angular velocity $\omega$ while the other is fixed we have

$$A = \pm \frac{a^2 b^2 \omega}{b^2 - a^2} \qquad \text{(36)},$$

where the upper sign is taken when the rotating cylinder is the outer one and the lower sign when the rotating cylinder is the inner one. Hence for both cases we have the critical angular velocity $\omega$ given by

$$\frac{a^2 b^2 \omega \rho}{(b^2 - a^2)\mu} = 4{\cdot}924 \qquad \text{(37)},$$

when $\lambda = 2$.

Thus we see that the stability of the steady motion is the same whether the outer cylinder is rotating or the inner cylinder is rotating if the angular velocities are equal in the two cases.

* It is observed that a disturbance of this type is likely to be set up if the axes of the two cylinders are parallel but not quite coincident.

## PART II.

### By W. J. HARRISON.

(I) Using two-dimensional polar coordinates $(r, \theta)$, let $V$ be the velocity in the steady motion of the liquid at any point between the cylinders $r=a$, and $r=b$, then $V=B/r+Cr$, where $B$ and $C$ are constants.

Let the velocity of the liquid in a disturbed motion be $(u, V+v)$. We obtain the discriminating equation by writing the right-hand side of equation (7) equal to zero. Thus

$$\mu\int_0^{2\pi}\int_a^b\left\{2\left(\frac{\partial u}{\partial r}\right)^2+2\left(\frac{1}{r}\frac{\partial v}{\partial\theta}+\frac{u}{r}\right)^2+\left(\frac{\partial v}{\partial r}+\frac{1}{r}\frac{\partial u}{\partial\theta}-\frac{v}{r}\right)^2\right\}rdrd\theta$$

$$=\int_0^{2\pi}\int_a^b\rho uv\left(\frac{V}{r}-\frac{\partial V}{\partial r}\right)rdrd\theta$$

$$=2B\rho\int_0^{2\pi}\int_a^b\frac{uv}{r}drd\theta.$$

If $\mu$ be greater than the value determined by this equation, the assumed type of disturbed motion must depend on the time in such a way that its kinetic energy will decrease.

(1) If $V=0$, $r=a$; $V=b\omega_0$, $r=b$; then $B=-a^2b^2\omega_0/(b^2-a^2)$.

(2) If $V=a\omega_1$, $r=a$; $V=0$, $r=b$; then $B=a^2b^2\omega_1/(b^2-a^2)$.

In both cases, $u$, $v$ being derivable from a stream function $\psi$, $\psi=0$ and $\dfrac{\partial\psi}{\partial r}=0$ at $r=a$, $r=b$.

It is clear that the critical value of $\mu$ is the same both for (1) and (2), if $\omega_0=\omega_1$, for a given disturbance. Hence the greatest value of $\mu$ for which instability is possible is the same in both cases, if $\omega_0=\omega_1$. This is the conclusion arrived at in Part I, where the maxima values are obtained, following the method of Orr.

To pursue the question further it is necessary to follow the method of Reynolds and assume a particular type of disturbance.

Let
$$u=-\frac{1}{r}\frac{\partial\psi}{\partial\theta},\quad v=\frac{\partial\psi}{\partial r},$$

where $\psi=f(r)\cos n\theta+F(r)\sin n\theta$, and $n$ is integral.

It is necessary that

$$f(a)=f(b)=F(a)=F(b)=0,$$
$$f'(a)=f'(b)=F'(a)=F'(b)=0.$$

The discriminating equation becomes

$$\mu\int_a^b\left[4n^2\{rf'(r)-f(r)\}^2+4n^2\{rF'(r)-F(r)\}^2+\{r^2f''(r)-rf'(r)+n^2f(r)\}^2\right.$$
$$\left.+\{r^2F''(r)-rF'(r)+n^2F(r)\}^2\right]\frac{dr}{r^3}$$

$$=2\beta\rho n\int_a^b\{F(r)f'(r)-F'(r)f(r)\}\frac{dr}{r^2}.$$

It is assumed in the first place that $f(r)$ is of the form $r^p(r-a)^2(r-b)^2$, and that $F(r)$ is of the form $r^q(r-a)^2(r-b)^2$. For purposes of numerical calculation $a$ and $b$ are taken to be equal to 1 and 2, respectively. The results of calculation are shown in the following table:

*Critical Value of $2B\rho/\mu$.*

| $n$ | $\psi=(r-2)^2(r-1)^2\{f_1(r)\cos n\theta+f_2(r)\sin n\theta\}$ | | | | | |
|---|---|---|---|---|---|---|
| | $f_1(r)=r$, $f_2(r)=1$ | $r$, $r^{-1}$ | $r$, $r^{-2}$ | $1$, $r^{-1}$ | $1$, $r^{-2}$ | $r^{-1}$, $r^{-2}$ |
| 1 | 6200 | 3800 | 3400 | 5200 | 3200 | 5000 |
| 2 | 3300 | 2000 | 1800 | 2800 | 1700 | 2700 |
| 3 | 2400 | 1500 | 1300 | 2100 | 1250 | 2000 |
| 4 | 2100 | 1300 | 1150 | 1800 | 1100 | 1800 |
| | A | B | C | D | E | F |

*Mean Critical Value of $2B\rho/\mu$.*

| Columns | $n=1$ | 2 | 3 | 4 |
|---|---|---|---|---|
| AB | 5000 | 2650 | 1950 | 1700 |
| ABCD | 4650 | 2500 | 1800 | 1600 |
| CDEF | 4200 | 2250 | 1650 | 1450 |
| EF | 4100 | 2200 | 1650 | 1450 |

It is probable that any disturbance set up would have a greater magnitude in the neighbourhood of the rotating cylinder than in the neighbourhood of the other. The calculations given above therefore tend to show that the motion when the outer cylinder is rotating is more stable than when the inner cylinder is in motion. This conclusion is confirmed by the following series of calculations.

It is assumed that

$$\psi=f(r)\cos n\theta+F(r)\sin n\theta,$$

where

$$f(r)=A\left(\frac{r-1}{c-1}\right)^2, \qquad 1<r<c,$$

$$f(r)=A\left(\frac{r-2}{c-2}\right)^2, \qquad c<r<2,$$

$$F(r)=\left(\frac{r-1}{d-1}\right)^2, \qquad 1<r<d,$$

$$F(r)=\left(\frac{r-2}{d-2}\right)^2, \qquad d<r<2,$$

and in each particular case $A$ is so chosen as to make the corresponding value of $2B\rho/\mu$ a minimum.

*Minimum Critical Values of $2B\rho/\mu$.*

$n = 1.$

| $d$ | $c = 1\cdot9$ | $1\cdot7$ | $1\cdot5$ | $1\cdot3$ | $1\cdot1$ |
|---|---|---|---|---|---|
| 1·9 | — | 12600 | 7000 | 9255 | 50200 |
| 1·7 | 12600 | — | 1330 | 1415 | 6675 |
| 1·5 | 7000 | 1330 | — | 740 | 3000 |
| 1·3 | 9255 | 1415 | 740 | — | 2870 |
| 1·1 | 50200 | 6675 | 3000 | 2870 | — |
| Average | 19800 | 5500 | 3000 | 3600 | 15700 |

$n = 2.$

| $d$ | $c = 1\cdot9$ | $1\cdot7$ | $1\cdot5$ | $1\cdot3$ | $1\cdot1$ |
|---|---|---|---|---|---|
| 1·9 | | 6840 | 4260 | 5940 | 25950 |
| 1·7 | 6840 | — | 860 | 980 | 3690 |
| 1·5 | 4260 | 860 | — | 570 | 1860 |
| 1·3 | 5940 | 980 | 570 | — | 1860 |
| 1·1 | 25950 | 3690 | 1860 | 1860 | — |
| Average | 10750 | 3100 | 1900 | 2350 | 8350 |

It is thus a fair presumption that the motion when the inner cylinder is rotating is less stable than when the outer cylinder is rotating, $\omega_0$ being equal to $\omega_1$.

As an example of a different type of disturbance, consider

$$f(r) = r^2 \sin^2 \left(\frac{p\pi(r-a)}{b-a}\right),$$

$$F(r) = r^2 \sin^2 \left(\frac{q\pi(r-a)}{b-a}\right),$$

where $p$ and $q$ are integers.

The minimum value of $2B\rho/\mu$, for integral values of $p$ and $q$, is approximately

$$\frac{34\pi^4 (b^6 - a^6)}{5(b-a)^4 (b^2 + a^2)}.$$

If $b = 2$, $a = 1$, this critical value is approximately $85\cdot68\pi^4$.

Even with these simple types of disturbance the labour of calculation is considerable. It will be observed that the theoretical minima critical velocities obtained in Part I are not approached. The very low limits obtained there for the critical angular velocity is evidence of a certain inherent instability. It is, accordingly, inferred from the experimental evidence, that the special types of disturbance corresponding to these limits are likely to be approximately set up when the inner cylinder is rotating, but not when the outer cylinder is rotating.

(II) It is stated by Lamb* that the motion in which the inner cylinder rotates and the outer is at rest is necessarily unstable, since other distributions of velocity than that in the accepted state of steady motion are possible which have less kinetic energy for the same angular momentum. The author can find no trace of any formal treatment of this point, so that some consideration of it may not be out of place. He has also been in communication with Prof. Lamb on the subject of this remark.

It is a simple matter to show that it is indifferent which cylinder is made to rotate; other distributions of velocity can be found which give less kinetic energy for the same angular momentum.

Let the velocity of the liquid be $v = f(r)$, where $r$ is the distance from the common axis of the cylinders.

The angular momentum $= A = \int_a^b 2\pi\rho r^2 f(r)\, dr$.

The kinetic energy $= T = \int_a^b \pi\rho r f^2(r)\, dr$.

Writing $r^2 = x$, $rf(r) = F(x)$, we have

$$A/\pi\rho = \int_{a^2}^{b^2} F(x)\, dx,$$

$$2T/\pi\rho = \int_{a^2}^{b^2} F^2(x)\, \frac{dx}{x}.$$

We wish to find the form of $F(x)$ so as to make $T$ a minimum for a given value of $A$, subject to

$$\begin{aligned} (1)\quad & F(x) = 0, & r = a, \\ & F(x) = bV, & r = b, \\ \text{or } (2)\quad & F(x) = aV, & r = a, \\ & F(x) = 0, & r = b. \end{aligned}$$

In the actual motion (1) $F(x) = \dfrac{b^2(x - a^2)V}{(b^2 - a^2)x^{\frac{1}{2}}}$.

(2) $F(x) = \dfrac{a^2(b^2 - x)V}{(b^2 - a^2)x^{\frac{1}{2}}}$.

Using the method of the Calculus of Variations, we must have

$$\int_{a^2}^{b^2} F(x)\, \eta \frac{dx}{x} = 0,$$

where $\eta$ is any function of $x$ satisfying $\int_{a^2}^{b^2} \eta\, dx = 0$.

* *Hydrodynamics*, 4th Ed., p. 655.

The solution is $F(x) = \lambda x$, where $\lambda$ is a constant which can be chosen so as to satisfy

$$A/\pi\rho = \int_{a^2}^{b^2} F(x)\,dx.$$

The boundary conditions cannot be satisfied using this form for $F(x)$, but by taking $F(x) = \lambda x$ we obtain the lower limit towards which the kinetic energy may be made to approach when the conditions are satisfied by taking a distribution of velocity which differs from that given by $F(x) = \lambda x$, only in the neighbourhood of $r = a$ and $r = b$.

In the actual motion (1)

$$\frac{A}{\pi\rho} = \frac{2b^2(b-a)(b+2a)V}{3(a+b)},$$

$$\frac{T}{2\pi\rho} = \left[\frac{b^2(a^2+b^2)}{b^2-a^2} - \frac{4a^2b^4}{(b^2-a^2)^2}\log(b/a)\right]V^2.$$

Taking $F(x) = \lambda x$, we find that

$$\int_{a^2}^{b^2} \lambda x\,dx = A/\pi\rho, \text{ if } \lambda = \frac{4b^2(b+2a)V}{3(a^2+b^2)(a+b)^2}.$$

With this value of $\lambda$

$$\int_{a^2}^{b^2} \lambda^2 x^2 \frac{dx}{x} = \frac{8b^4(b+2a)^2(b-a)V^2}{9(a^2+b^2)(a+b)^3}$$

$$= T'/2\pi\rho,$$

where $T'$ is the kinetic energy in the hypothetical motion.

The relative values of $T$ and $T'$ are indicated by the data given in the following table

| $b$ = | $3a$ | $2a$ | $1{\cdot}3a$ | $1{\cdot}1a$ |
|---|---|---|---|---|
| $100(T-T')/T =$ | 1 | 3 | 7 | 17 |

In the actual motion (2)

$$\frac{A}{\pi\rho} = \frac{2a^2(2b+a)(b-a)V}{3(a+b)},$$

$$\frac{T}{2\pi\rho} = \left[\frac{a^2(a^2+b^2)}{b^2-a^2} - \frac{4a^4b^2}{(b^2-a^2)^2}\log(b/a)\right]V^2.$$

In this case

$$\lambda = \frac{4a^2(2b+a)V}{3(a+b)^2(a^2+b^2)},$$

$$\frac{T'}{2\pi\rho} = \frac{8a^4(2b+a)^2(b-a)V^2}{9(a^2+b^2)(a+b)^3}.$$

The relation between $T$ and $T'$ is given below:

| $b$ = | $3a$ | $2a$ | $1{\cdot}3a$ | $1{\cdot}1a$ |
|---|---|---|---|---|
| $100(T-T')/T$ | 78 | 63 | 39 | 30 |

If the difference between $T$ and $T'$ be accepted as a criterion of instability, the motion when the inner cylinder is rotated is much more unstable than when the outer cylinder is rotated, and the tendency to instability increases as the distance between the cylinders is increased. On the other hand when the outer cylinder is rotated the tendency to instability should increase as this distance is diminished, which is contrary to usual experience.

This criterion is independent of the degree of viscosity of the liquid, and of the angular velocity of the cylinder, and therefore falls into an altogether different category from the accepted criteria for other modes of motion.

(III) Reynolds, in the course of his investigations, obtains the equations which determine the effect of the turbulent motion on the mean-mean motion which, in consequence of the existence of the disturbance, differs from the steady motion. But, if reference be made to equation (64) on page 572 of Reynolds' *Scientific Papers*, vol. II, it will be seen that he finally neglects terms of the fourth degree in the velocities of the relative-mean motion, thereby reducing his work in the end to a consideration of small disturbances only, and consequently his results differ in no respect from those of other investigators who make this simplifying assumption from the start. His criterion, therefore, refers to incipient turbulent motion only.

If the neglected terms of the fourth degree be retained in the discriminating equation, the condition that the kinetic energy of a given type of initial disturbance is stationary becomes of the form $AU_c = B\mu + C/\mu$, instead of $AU_c = B\mu$, where, in the case of flow between parallel planes, $U_c$ is the mean undisturbed velocity, $A$, $B$ depend on the square of the velocity of the relative mean motion, and $C$ depends on the fourth power.

The effect of the additional term $C/\mu$ is easily seen. For given values of $\mu$ and $U_c$, an upper limit is set to the amplitude of a given type of initial disturbance in order that the kinetic energy of the disturbance may increase. If the amplitude is greater than this limit the kinetic energy must decrease. Further this limit may be made as small as is desired by sufficiently decreasing $\mu$.

It would therefore appear that the results of Reynolds' investigations are brought into agreement with Rayleigh's conclusions for a non-viscous liquid, namely that the steady motion of an inviscid liquid between two parallel planes is stable subject to the condition that $\frac{d^2U}{dz^2}$ is one-signed, where $U = f(z)$ gives the distribution of velocity in the steady motion and the boundaries are parallel planes perpendicular to the axis of $z$.

This question has been discussed by G. I. Taylor* with the result that Rayleigh's conclusions are verified by means of entirely different considerations. At the same time Taylor maintains that there may be a finite difference in behaviour between a perfectly inviscid liquid and one which has an infinitesimal viscosity. It would appear from the considerations given above that this need not be inferred from the discrepancy discussed in this section, since this discrepancy is apparent only, and arises from a premature approximation.

* *Phil. Trans. Roy. Soc.*, vol. 215 A, pp. 23—26, 1915.

# XXIII. *On a General Infinitesimal Geometry, in reference to the Theory of Relativity.*

By Wilhelm Wirtinger (Vienna).

[*Received* July 23, *Read* October 31, 1921.]

The purpose of this paper is to study some ideas suggested by the infinitesimal geometry of Weyl*, especially his *Parallelverschiebung*, in the most general form, as I believe, in accordance with present-day needs. The form adopted may be too general for immediate application to Physics; it is general enough to provide for a great variety of possibilities of experience in Physics and Astronomy.

1. We take $n$ variables, $x_\alpha\,(\alpha=1, 2, \ldots, n)$, to which we apply all *point-transformations* $x_\alpha=\phi_\alpha(x_\beta')$, whose Jacobian does not vanish. We take a second system of variables, $\xi^\alpha\,(\alpha=1,2,\ldots,n)$, transformed like the differentials $dx_\alpha$. A third system of variables, denoted by $u_\alpha$, is also used, transformed like the differential coefficients $\partial\phi/\partial x_\alpha$, where $\phi$ is any function of the $x_\alpha$. Thus, denoting the transformed variables by an accent, we have

$$\xi'^\alpha=\frac{\partial x_\alpha'}{\partial x_\beta}\xi^\beta,\quad u_\alpha'=\frac{\partial x_\beta}{\partial x_\alpha'}u_\beta \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots(1),$$

wherein, after Einstein, the sign of summation is omitted; this is in regard to repeated indices.

We consider one or more systems of variables like the $\xi^\alpha$, say $\eta^\alpha$, $\zeta^\alpha$, and one or more systems like the $u_\alpha$, say, $v_\alpha$, $w_\alpha$. The group of point-transformations of the $x_\alpha$, mentioned above, and its extension to the $\xi^\alpha$, $u_\alpha$ given by (1), we call $P$.

The $x_\alpha$ being varied infinitesimally (or differentiated with regard to a parameter), the variations are like the $\xi^\alpha$, and may be so denoted. We consider corresponding differentials of the $\xi^\alpha$ and $u_\alpha$, say, rather, of the $\eta^\alpha$ and $v_\alpha$, denoted by $\delta_\xi\eta^\alpha$, $\delta_\xi v_\alpha$, whose transformed values, in accordance with (1), are given by

$$\delta_{\xi'}\eta^{\alpha'}=\frac{\partial x_\alpha'}{\partial x_\beta}\delta_\xi\eta^\beta+\frac{\partial^2 x_\alpha'}{\partial x_\beta\partial x_\gamma}\xi^\gamma\eta^\beta \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots(2),$$

$$\delta_{\xi'}v_\alpha'=\frac{\partial x_\beta}{\partial x_\alpha'}\delta_\xi v_\beta+\frac{\partial^2 x_\beta}{\partial x_\alpha'\partial x_\gamma'}\xi'^\gamma v_\beta.$$

2. If a new system of variables be such that the $\delta_{\xi'}\eta'$, $\delta_{\xi'}v'$ are zero, we must have

$$\delta_\xi\eta^\alpha=-\frac{\partial x_\alpha}{\partial x_\beta'}\frac{\partial^2 x_\beta'}{\partial x_\gamma\partial x_\delta}\xi^\delta\eta^\gamma \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots(3),$$

$$\delta_\xi v_\alpha=-\frac{\partial x_\beta'}{\partial x_\alpha}\frac{\partial^2 x_\gamma}{\partial x_\beta'\partial x_\delta'}\xi'^\delta v_\gamma.$$

Thus the $\delta_\xi\eta^\alpha$ are symmetrically bilinear in the $\xi$, $\eta$, and we may put

$$\delta_\xi\eta^\alpha=a^\alpha_{\beta\gamma}\eta^\beta\xi^\gamma \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots(4).$$

From the identity

$$\frac{\partial}{\partial x_\delta}\left(\frac{\partial x_\alpha}{\partial x_\beta'}\cdot\frac{\partial x_\beta'}{\partial x_\gamma}\right)=\frac{\partial^2 x_\alpha}{\partial x_\beta'\partial x_\lambda'}\frac{\partial x_\lambda'}{\partial x_\delta}\frac{\partial x_\beta'}{\partial x_\gamma}+\frac{\partial x_\alpha}{\partial x_\beta'}\frac{\partial^2 x_\beta'}{\partial x_\gamma\partial x_\delta}=0,$$

* H. Weyl, *Mathematische Zeitschrift*, II. (1918), pp. 394 ff.

we then have $$\delta_\xi v_\alpha = -a^\beta_{\gamma\alpha} v_\beta \xi^\gamma \qquad (4\,a),$$

with $$a^\alpha_{\beta\gamma} = -\frac{\partial x_\alpha}{\partial x_\delta'} \cdot \frac{\partial^2 x_\delta'}{\partial x_\beta \partial x_\gamma},$$

or $$\frac{\partial^2 x_\alpha'}{\partial x_\beta \partial x_\gamma} = -a^\epsilon_{\beta\gamma} \frac{\partial x_\alpha'}{\partial x_\epsilon} \qquad (5),$$

for determination of the new variables $x_\alpha'$; and the conditions of integrability

$$K^\alpha_{\beta\gamma\delta} = \frac{\partial a^\alpha_{\beta\gamma}}{\partial x_\delta} - \frac{\partial a^\alpha_{\beta\delta}}{\partial x_\gamma} + a^\mu_{\beta\gamma} a^\alpha_{\delta\mu} - a^\mu_{\beta\delta} a^\alpha_{\gamma\mu} = 0 \qquad (6).$$

If the $a^\alpha_{\beta\gamma}$ satisfy these conditions, the $x_\alpha'$, defined by (5), transform the $\delta_\xi \eta^\alpha$, $\delta_\xi v_\alpha$ to zero; all systems $x_\alpha'$ of this kind are transformed together by an ordinary linear substitution with constant coefficients, and are therefore *affin* together. It is remarkable that (6) is only a consequence of the behaviour of the $\delta_\xi \eta$. $\delta_\xi v$ under the transformation $P$, without any other assumption for these variables.

We may also introduce the $\delta_\xi \eta$, $\delta_\xi v$ by extension (*Erweiterung*) of $P$, in the notation of S. Lie; for this we should consider the $x_\alpha$ as functions of two parameters, say $t, \tau$, and then consider the transformations of the $\partial^2 x_\alpha/\partial t \partial\tau$, like the $\delta_\xi \eta^\alpha$, and of $\partial(\partial\phi/\partial x_\alpha)/\partial\tau$, like the $\delta_\xi v_\alpha$.

3. A more intuitive, or geometric, interpretation is obtained by regarding the $x_\alpha$ as Gauss parameters on a manifoldness $M_n$, of $n$ dimensions, in a linear space of $m\,(> n)$ dimensions. There is at any point, $O$, of $M_n$ a linear tangential manifoldness, $E_n$, in which is a bundle of rays touching $M_n$ at this point: in $E_n$ is also a bundle of $E_{n-1}$ passing through this point. The $\eta^\alpha$ may be taken for homogeneous coordinates of the rays, and the $v_\alpha$ for the dualistic coordinates of the $E_{n-1}$. An infinitesimal dislocation of the point of $M_n$ varies the $\eta^\alpha$ and $v_\alpha$; the infinitesimally neighbouring elements have the coordinates $\eta^\alpha + \delta_\xi \eta^\alpha$, $v_\alpha + \delta_\xi v_\alpha$.

Now let us imagine an observer at the point $O$ of $M_n$, who has a memory. Along the different $\eta^\alpha$, $v_\alpha$ come into his mind the "events," or some indications of the events, in the world of the $x_\alpha$. In the space of the $\eta^\alpha$, $v_\alpha$, that is in the bundle of which $O$ is the vertex, he can and will establish a projective geometry, and order his impressions in accord with this; this geometry will apparently be independent of the special system of $x_\alpha$. If his position on the $M_n$ is (infinitesimally) changed, his system of $\eta^\alpha$, $v_\alpha$ is varied; if the $\delta\eta$, $\delta v$ are in agreement with the conditions (5), (6), he is able to describe that variation, which is all he can observe, by a particular system of $x_\alpha$, in which the $\delta\eta^\alpha$, $\delta v_\alpha$ all vanish. If he describes a curve in the world of the $x_\alpha$, his impressions may be described by representing the single $\eta^\alpha$, $v_\alpha$ as functions of a single parameter; if a group of events shews the properties of a transformation-group of one parameter, he will put this group into the normal form, where the parameter is additive. If we take $n = 4$, we have a three-dimensional intuitive space with projective geometry, and an additive parameter like our old-fashioned notion of time. The parameter $t$, mentioned before, may be derived from the apparent motion of the fixed stars in the sky. It is remarkable that there is no need for the assumption of a metric geometry. We may call the theory of the events in the $(\eta^\alpha, v_\alpha)$-space, subjective physics, thus distinguishing from objective physics in the world of the $x_\alpha$. In mathematical terms, perhaps the weightiest reason for a real objective world is the possibility of describing in terms of some $x_\alpha$ the phenomena subjectively observed in the $\eta^\alpha$, $v_\alpha$, which are the only phenomena we are able to receive.

4. If we accept this point of view, there is an interest in developing a hypothesis for the $\delta_\xi \eta^a$, $\delta_\xi v_a$ ample enough for a reasonable theory, but capable of analytical treatment. We may consider in the $(\eta^a, v_a)$-bundle, with $O$ as centre, the $E_{n-1}$-elements consisting of a ray, of coordinates $\eta^a$, with an incident $E_{n-1}$ of coordinates $v_a$, so that $\eta^a v_a = 0$. These elements form a $M_{2n-3}$. If $O$ is moved, the $E_{n-1}$-elements may remain elements of the same kind, undergoing a *contact-transformation*, in the sense of S. Lie.

Denote the differentials of $\eta^a$, $v_a$, in the $O$-bundle, by $d\eta^a$, $dv_a$; and the variation caused by the variation of $x_a$, by $\delta\eta^a$, $\delta v_a$. The signs $d$ and $\delta$ are commutable, and the conditions for a *contact-transformation* are

$$\delta_\xi(\eta^a dv_a) = \delta_\xi \eta^a dv_a + \eta^a \delta_\xi dv_a = \rho \eta^a dv_a + \sigma d\eta^a v_a \quad \text{.....................(7);}$$

but, from

$$\eta^a dv_a + d\eta^a v_a = 0 \quad \text{.....................(7a),}$$

we have

$$\delta_\xi v_a d\eta^a + v_a \delta_\xi d\eta^a = -\rho\eta^a dv_a - \sigma d\eta^a v_a \quad \text{.....................(7b),}$$

and we have

$$\delta_\xi \eta^a v_a + \eta^a \delta_\xi v_a = 0 \quad \text{.....................(7c).}$$

The $\delta_\xi \eta^a$, $\delta_\xi v_a$ are linear homogeneous functions of the $\xi^a$, as are the $\rho$ and $\sigma$. The $\delta_\xi \eta^a$ are homogeneous of order one, not necessarily linear, in the $\eta^a$; and are of order zero in the $v_a$. *Vice versa* the $\delta_\xi v_a$ are homogeneous of order one in the $v_a$ and of order zero in the $\eta^a$. But $\rho$, $\sigma$ are homogeneous of order zero in both kinds of variables $\eta$, $v$. All this is in accord with the transformation formulae (2).

Following Lie we put

$$\eta^a \delta_\xi v_a = W(x, \eta, v, \xi) \quad \text{.....................(8),}$$

so that $W$ is homogeneous and linear in the $\xi^a$, and of order one in $\eta^a$, $v_a$. By differentiation we then have

$$\frac{\partial W(x, \eta, v, \xi)}{\partial \eta^\beta} = \delta_\xi v_\beta + \eta^a \frac{\partial \delta_\xi v_a}{\partial \eta^a},$$

$$\frac{\partial W(x, \eta, v, \xi)}{\partial v_\beta} = \eta^a \frac{\partial \delta_\xi v_a}{\partial v_\beta} \quad \text{.....................(9);}$$

using the equations

$$\delta_\xi d\eta^a = \frac{\partial \delta_\xi \eta^a}{\partial \eta^\beta} d\eta^\beta + \frac{\partial \delta_\xi \eta^a}{\partial v_\beta} dv_\beta,$$

$$\delta_\xi dv_a = \frac{\partial \delta_\xi v_a}{\partial \eta^\beta} d\eta^\beta + \frac{\partial \delta_\xi v_a}{\partial v_\beta} dv_\beta,$$

substituting in (7), and comparing coefficients of $d\eta^a$, $dv_a$, we find

$$\delta_\xi \eta^a = -\frac{\partial W(x, \eta, v, \xi)}{\partial v_a} + \rho\eta^a,$$

$$\delta_\xi v_a = \frac{\partial W(x, \eta, v, \xi)}{\partial \eta^a} - \sigma v_a \quad \text{.....................(9 a);}$$

if these be used in (7), (7 *a*), Euler's theorem of homogeneous functions shews that there is no other condition for the functions $W$, $\rho$, $\sigma$.

Without essential modification, we may put $W + \mu\eta^a v_a$ in place of $W$, and so make $\rho = \sigma$. Then $W$ is unique. And finally we have

$$\delta_\xi \eta^a = -\frac{\partial W(x, \eta, v, \xi)}{\partial v_a} + \rho(x, \eta, v, \xi)\eta^a,$$

$$\delta_\xi v_a = \frac{\partial W(x, \eta, v, \xi)}{\partial \eta^a} - \rho(x, \eta, v, \xi) v_a \quad \text{.....................(10).}$$

The formulae (10) satisfy the most general assumption for a $\delta$-transformation effecting a relation between the infinitesimally near bundles $(O)$ and $(O')$, subject to the contact condition. In particular, partial differential equations, and their solutions, remain such.

By a transformation of the group $P$ we have

$$W'(x', \eta', v', \xi') = W(x, \eta, v, \xi) - \frac{\partial^2 x_a'}{\partial x_\beta \partial x_\gamma} v_a' \eta^\beta \xi^\gamma,$$

$$\rho'(x', \eta', v', \xi') = \rho(x, \eta, v, \xi) \quad \ldots\ldots (11);$$

thus $W$ can only be transformed to zero if it is linear in $\eta$, $v$, $\xi$ and symmetrical in regard to $\eta$ and $\xi$; the equations (6) then give the conditions for the vanishing of $\delta_\xi \eta^a$, $\delta_\xi v_a$ in a particular system of coordinates.

5. We consider now homogeneous, but not necessarily rational functions, of one or more series of variables $\eta$, $v$; $\zeta$, $w$, also depending upon the $x_a$; denoting such a function by $T(x, \eta, v, \zeta, w, \ldots)$. On account of $\eta^a v_a = 0$, $\zeta^a w_a = 0$, etc., the operator

$$\Theta T = \frac{\partial^2 T}{\partial \eta^a \partial v_a} \quad \ldots\ldots (12)$$

is an invariant, and its repetition leads to forms invariant for the $P$ transformation. A canonical form can be reached by applying the well-known Clebsch-Gordan expansion. The *contraction* (*Verjüngung*), of the usual tensor-analysis, is a special case of this process. The process can only be repeated a finite number of times when $T$ is an integral rational function. This was remarked by F. Klein.

We consider the behaviour of $T$ caused by a displacement of $O$, after (10), given by

$$\delta_\xi T = \frac{\partial T}{\partial x_a} \xi^a - \frac{\partial T}{\partial \eta^\beta} \frac{\partial W(x, \eta, v, \xi)}{\partial v_\beta} + \frac{\partial T}{\partial v_\beta} \frac{\partial W(x, \eta, v, \xi)}{\partial \eta^\beta} + \rho(x, \eta, v, \xi)(m - \mu) T \quad \ldots\ldots (13),$$

$m$ denoting the degree in $\eta$, and $\mu$ in $v$.

The vanishing of $\delta_\xi T$, for every $\xi$, is the condition for the invariance of $T$ by every displacement. With a given function $W$, which only affects the proportions of the $\eta^a$, $v_a$, we can, by choosing $\rho$ conveniently, make any function $T$ invariant; that means an adaptation of the direct measures of the $\eta^a$, $v_a$ to the transformation $W$. The geometry of Weyl is obtained if for $T$ we take the usual $ds^2$, with $\rho$ independent of $\eta$, $v$. This measurement cannot be applied to the elements of $T = 0$, in perfect accord with the behaviour of the usual measurement for $ds^2 = 0$.

6. We now combine a displacement $\xi^a$ with another one, $\zeta^a$, and find

$$\delta_\zeta \delta_\xi T - \delta_\xi \delta_\zeta T = \frac{\partial T}{\partial x_a}(\delta_\zeta \xi^a - \delta_\xi \zeta^a) + \frac{\partial T}{\partial \eta^\beta}(\delta_\zeta \delta_\xi \eta^\beta - \delta_\xi \delta_\zeta \eta^\beta) + \frac{\partial T}{\partial v_\beta}(\delta_\zeta \delta_\xi v_\beta - \delta_\xi \delta_\zeta v_\beta) \quad \ldots\ldots (14),$$

which is evidently an invariant in regard to $P$. After some easy calculations we find, for the behaviour of the single terms on the right side, under $P$, the equations

$$(\delta_\zeta \delta_\xi \eta^a - \delta_\xi \delta_\zeta \eta^a)' = \frac{\partial x_a'}{\partial x_\beta}(\delta_\zeta \delta_\xi \eta^\beta - \delta_\xi \delta_\zeta \eta^\beta) + \frac{\partial^2 x_a'}{\partial x_\beta \partial x_\gamma}(\delta_\zeta \xi^\gamma - \delta_\xi \zeta^\gamma),$$

$$(\delta_\zeta \delta_\xi v_a - \delta_\xi \delta_\zeta v_a)' = \frac{\partial x_\beta}{\partial x_a'}(\delta_\zeta \delta_\xi v_\beta - \delta_\xi \delta_\zeta v_\beta) + \frac{\partial^2 x_\beta}{\partial x_a' \partial x_\gamma'} \frac{\partial x_\gamma'}{\partial x_\gamma} v_\gamma (\delta_\zeta \xi^\delta - \delta_\xi \zeta^\delta),$$

$$(\delta_\zeta \xi^a - \delta_\xi \zeta^a)' = \frac{\partial x_a'}{\partial x_\beta}(\delta_\zeta \xi^\beta - \delta_\xi \zeta^\beta) \quad \ldots\ldots (15);$$

with these we can establish the invariance of (14) by direct computation.

7. The expression in (14) is in a very close relation to the behaviour of $T$, when $O$ describes a small closed path back to its original position. To develop this relation we write more explicitly

$$\delta_\xi \eta^\alpha = Z_\beta{}^\alpha (x, \eta, v)\, \xi^\beta, \quad \delta_\xi v_\alpha = V_{\alpha\beta} (x, \eta, v)\, \xi^\beta \quad \text{.....................(16)},$$

and put

$$x_\alpha = x_\alpha{}^0 + r\psi_\alpha(s),$$

where the $\psi_\alpha(s)$ are periodic functions of $s$, with the period 1, and $\psi_\alpha(0)=0$, and $r$ is a real variable confined to a sufficiently small interval containing zero. When $s$ varies from 0 to 1, and $r$ is small enough, the $x_\alpha$ describe a small closed path from and back to $x_\alpha{}^0$. For the $\eta^\alpha$ and $v_\alpha$ we have the ordinary system of differential equations

$$\frac{d\eta^\alpha}{ds} = rZ_\beta{}^\alpha (x, \eta, v)\, \psi_\beta'(s), \quad \frac{dv_\alpha}{ds} = r V_{\alpha\beta} (x, \eta, v)\, \psi_\beta'(s) \quad \text{..................(17)};$$

the $Z, V, \psi$ being supposed regular in a small, but finite, neighbourhood of $\overset{0}{\eta}{}^\alpha$, $v_\alpha{}^0$, $x_\alpha{}^0$ and 0, we can develop the $\eta, v$ in series of powers of the parameter $r$. If

$$\eta^\alpha = \overset{0}{\eta}{}^\alpha + r\eta_1{}^\alpha + r^2\eta_2{}^\alpha + r^3\eta_3{}^\alpha + \dots,$$

$$v_\alpha = \overset{0}{v}_\alpha + rv_{1\alpha} + r^2 v_{2\alpha} + r^3 v_{3\alpha} + \dots,$$

we find, comparing the two sides of the differential equations, the recurrence-formulae

$$\frac{d\eta_p{}^\alpha}{ds} = \left(\frac{d^{p-1} Z_\beta{}^\alpha (x, \eta, v)}{dr^{p-1}}\right)_{r=0} \psi_\beta'(s), \quad \frac{dv_{p\alpha}}{ds} = \left(\frac{d^{p-1} V_{\alpha\beta} (x, \eta, v)}{dr^{p-1}}\right)_{r=0} \psi_\beta'(s) \quad \text{.........(18)},$$

which shew that we can compute the $\eta_p{}^\alpha$, $v_{p\alpha}$, step by step, by simple quadratures.

On account of the initial conditions $\eta_k{}^\alpha(0) = v_{\alpha k}(0) = 0$, for every $k$; thus

$$\eta_1{}^\alpha(s) = Z_\beta{}^\alpha (x^0, \eta^0, v^0)\, \psi_\beta(s), \quad v_{1\alpha}(s) = V_{\alpha\beta}(x^0, \eta^0, v^0)\, \psi_\beta(s) \quad \text{..................(19)};$$

from these, by substitution and integration, we find

$$\eta_2{}^\alpha(s) = \left(\frac{\partial Z_\beta{}^\alpha}{\partial x_\delta} + \frac{\partial Z_\beta{}^\alpha}{\partial \eta^\gamma} Z_\delta{}^\gamma + \frac{\partial Z_\beta{}^\alpha}{\partial v_\gamma} V_{\gamma\delta}\right) \int_0^s \psi_\delta \psi_\beta'\, ds,$$

$$v_{2\alpha}(s) = \left(\frac{\partial V_{\alpha\beta}}{\partial x_\delta} + \frac{\partial V_{\alpha\beta}}{\partial \eta^\gamma} Z_\delta{}^\gamma + \frac{\partial V_{\alpha\beta}}{\partial v_\gamma} V_{\gamma\delta}\right) \int_0^s \psi_\delta \psi_\beta'\, ds \quad \text{.....................(20)}.$$

Thus, if we put

$$\omega^{\delta\beta} = r^2 \int_0^1 \psi_\delta \psi_\beta'\, ds,$$

we have $\omega^{\delta\beta} = -\omega^{\beta\delta}$, and the first terms in our series give the well-known expressions whose vanishing are the conditions of complete integrability, which is instructive, (16) being general differential equations.

We may, further, suppose the $x_\alpha$ to be functions of the parameters, $\tau_1$, $\tau_2$, such that they describe an element of surface bounded by the path of integration, and then consider the limit of the quotients of the first terms by the integral $\int d\tau_1 d\tau_2$, extended over this element. Writing $x_\alpha$, $\eta^\alpha$, etc., for $x_\alpha{}^0$, $\overset{0}{\eta}{}^\alpha$, etc., these limits are

$$\frac{1}{2}\left(\frac{\partial Z_\beta{}^\alpha}{\partial x_\delta} - \frac{\partial Z_\delta{}^\alpha}{\partial x_\beta} + \frac{\partial Z_\beta{}^\alpha}{\partial \eta^\gamma} Z_\delta{}^\gamma - \frac{\partial Z_\delta{}^\alpha}{\partial \eta^\gamma} Z_\beta{}^\gamma + \frac{\partial Z_\beta{}^\alpha}{\partial v_\gamma} V_{\gamma\delta} - \frac{\partial Z_\delta{}^\alpha}{\partial v_\gamma} V_{\gamma\beta}\right) \frac{\partial (x_\delta, x_\beta)}{\partial (\tau_1, \tau_2)},$$

and

$$\frac{1}{2}\left(\frac{\partial V_{\alpha\beta}}{\partial x_\delta} - \frac{\partial V_{\alpha\delta}}{\partial x_\beta} + \frac{\partial V_{\alpha\beta}}{\partial \eta^\gamma} Z_\delta{}^\gamma - \frac{\partial V_{\alpha\delta}}{\partial \eta^\gamma} Z_\beta{}^\gamma + \frac{\partial V_{\alpha\beta}}{\partial v_\gamma} V_{\gamma\delta} - \frac{\partial V_{\alpha\delta}}{\partial v_\gamma} V_{\gamma\beta}\right) \frac{\partial (x_\delta, x_\beta)}{\partial (\tau_1, \tau_2)},$$

which are evidently transformed by the group $P$ respectively like $\eta^\alpha$ and $v_\alpha$.

As the $\partial x_\alpha/\partial\tau_1, \partial x_\alpha/\partial\tau_2$ are transformed like the $\eta^\alpha$, we may write for them $\xi^\alpha$, $\zeta^\alpha$; then, introducing two new variables $\phi^\alpha$, $f_\alpha$, the former like $\eta^\alpha$, the second like $v_\alpha$, we infer that

$$\frac{1}{2}\left(\frac{\partial Z_\beta{}^\alpha}{\partial x_\delta}-\frac{\partial Z_\delta{}^\alpha}{\partial x_\beta}+\frac{\partial Z_\beta{}^\alpha}{\partial\eta^\gamma}Z_\delta{}^\gamma-\frac{\partial Z_\delta{}^\alpha}{\partial\eta^\gamma}Z_\beta{}^\gamma+\frac{\partial Z_\beta{}^\alpha}{\partial v_\gamma}V_{\gamma\delta}-\frac{\partial Z_\delta{}^\alpha}{\partial v_\gamma}V_{\gamma\beta}\right)f_\alpha(\xi^\delta\zeta^\beta-\xi^\beta\zeta^\delta)$$
$$= M^\alpha_{\delta\beta}f_\alpha(\xi^\delta\zeta^\beta-\xi^\beta\zeta^\delta)=M(x,\eta,v;\ \xi,\zeta,f) \quad\text{.................(21}a\text{)},$$

and
$$\frac{1}{2}\left(\frac{\partial V_{\alpha\beta}}{\partial x_\delta}-\frac{\partial V_{\alpha\delta}}{\partial x_\beta}+\frac{\partial V_{\alpha\beta}}{\partial\eta^\gamma}Z_\delta{}^\gamma-\frac{\partial V_{\alpha\delta}}{\partial\eta^\gamma}Z_\beta{}^\gamma+\frac{\partial V_{\alpha\beta}}{\partial v_\gamma}V_{\gamma\delta}-\frac{\partial V_{\alpha\delta}}{\partial v_\gamma}V_{\gamma\beta}\right)\phi^\alpha(\xi^\delta\zeta^\beta-\xi^\beta\zeta^\delta)$$
$$= N_{\alpha\delta\beta}\phi^\alpha(\xi^\delta\zeta^\beta-\xi^\beta\zeta^\delta)=N(x,\eta,v;\ \xi,\zeta,\phi) \quad\text{..............(21}b\text{)},$$

are both invariant under the transformations $P$. These expressions $M$ and $N$ are the immediate generalisations of the well-known Riemann-Christoffel curvature tensor.

8. We may also remark that, when the $x_\alpha$ describe such a small closed path, any form $T$ undergoes the variation

$$\frac{\partial T}{\partial\eta^\alpha}M^\alpha_{\delta\beta}(\xi^\delta\zeta^\beta-\xi^\beta\zeta^\delta)+\frac{\partial T}{\partial v_\alpha}N_{\alpha\delta\beta}(\xi^\delta\zeta^\beta-\xi^\beta\zeta^\delta) \quad\text{........................(22)},$$

which is also invariant under $P$.

Referring to formulae (15), the $M$, $N$ are

$$M=(\delta_\xi\delta_\zeta\eta^\alpha-\delta_\zeta\delta_\xi\eta^\alpha)f_\alpha-Z_\beta{}^\alpha f_\alpha(\delta_\zeta\xi^\beta-\delta_\xi\zeta^\beta),$$
$$N=(\delta_\xi\delta_\zeta v_\alpha-\delta_\zeta\delta_\xi v_\alpha)\phi^\alpha-V_{\alpha\beta}\phi^\alpha(\delta_\zeta\xi^\beta-\delta_\xi\zeta^\beta) \quad\text{........................(23)},$$

and we may also remark the generalisation of the usual covariant differentiation

$$\frac{\partial T}{\partial x_\alpha}\xi^\alpha+\frac{\partial T}{\partial\eta^\beta}\delta_\xi\eta^\beta+\frac{\partial T}{\partial v_\beta}\delta_\xi v_\beta \quad\text{....................................(24)},$$

and an analogous formula for several pairs of variables like $\eta^\alpha$, $v_\alpha$. If the $M$, $N$ are identically zero, the system of total differential equations

$$d\eta^\alpha=Z_\beta{}^\alpha dx_\beta,\quad dv_\alpha=V_{\alpha\beta}dx_\beta \quad\text{................................(25)},$$

is completely integrable, and the $\eta^\alpha$, $v_\alpha$ may be represented in the forms

$$\eta^\alpha=\phi^\alpha(x,x^*,\eta^*,v^*),\quad v_\alpha=f_\alpha(x,x^*,\eta^*,v^*) \quad\text{........................(26)},$$

so that the rays, and the $E_{n-1}$ of an $O^*$-bundle, at $x_\alpha{}^*$, change into the rays and the $E_{n-1}$ of an $O$-bundle at $x_\alpha$, independently of the path described, in the manner of the parallel-displacement of ordinary Euclidian geometry.

If we now suppose that only the ratios of the $\eta^\alpha$, $v_\alpha$ are independent of the path, the meaning of the $M$, $N$ shews directly, and computation confirms, that

$$M^\alpha_{\delta\beta}=\eta^\alpha a_{\delta\beta}(x,\eta,v),\quad N_{\alpha\delta\beta}=v_\alpha b_{\delta\beta}(x,\eta,v),$$

wherein the $a$, $b$ are homogeneous functions of $\eta^\alpha$, $v_\alpha$ of zero degree.

With the operator arising from (24)

$$(\ )_\alpha=\frac{\partial}{\partial x_\alpha}+Z_\alpha{}^\beta\frac{\partial}{\partial\eta^\beta}+V_{\alpha\beta}\frac{\partial}{\partial v_\beta} \quad\text{................................(27)},$$

we have
$$(M^\alpha_{\delta\beta})_\epsilon+(M^\alpha_{\beta\epsilon})_\delta+(M^\alpha_{\epsilon\delta})_\beta=\frac{\partial Z_\beta{}^\alpha}{\partial\eta^\gamma}M^\gamma_{\epsilon\delta}+\frac{\partial Z_\epsilon{}^\alpha}{\partial\eta^\gamma}M^\gamma_{\delta\beta}+\frac{\partial Z_\delta{}^\alpha}{\partial\eta^\gamma}M^\gamma_{\beta\epsilon}$$
$$+\frac{\partial Z_\beta{}^\alpha}{\partial v_\gamma}N_{\gamma\epsilon\delta}+\frac{\partial Z_\epsilon{}^\alpha}{\partial v_\gamma}N_{\gamma\delta\beta}+\frac{\partial Z_\delta{}^\alpha}{\partial v_\gamma}N_{\gamma\beta\epsilon} \quad\text{........................(28)},$$

with an analogous formula obtained by changing $M$, $Z$ into $N$, $V$, respectively.

9. The formulae of §§ 7, 8 do not assume that the $\delta_\xi \eta^a$, $\delta_\xi v_a$ belong to a contact-transformation, and are quite general. If we assume this, and put

$$W(x, \eta, v, \xi) = W_a(x, \eta, v)\,\xi^a, \quad \rho(x, \eta, v, \xi) = \rho_a(x, \eta, v)\,\xi^a \quad \ldots\ldots\ldots\ldots(29),$$

we obtain

$$Z_\beta{}^a = -\frac{\partial W_\beta}{\partial v_a} + \rho_\beta \eta^a, \quad V_{a\beta} = \frac{\partial W_\beta}{\partial \eta^a} - \rho_\beta v_a \ldots\ldots\ldots\ldots(30),$$

and

$$M^a = \omega^{\delta\beta}\left(-\frac{\partial \Omega_{\delta\beta}}{\partial v_\beta} + \rho_{\delta\beta}\eta^a\right), \quad N_a = \omega^{\delta\beta}\left(\frac{\partial \Omega_{\delta\beta}}{\partial \eta^a} - \rho_{\delta\beta} v_a\right) \ldots\ldots\ldots\ldots(31),$$

where

$$\Omega_{\delta\beta} = \frac{\partial W_\beta}{\partial x_\delta} - \frac{\partial W_\delta}{\partial x_\beta} + \frac{\partial(W_\beta, W_\delta)}{\partial(v_\gamma, \eta^\gamma)},$$

$$\rho_{\delta\beta} = \frac{\partial \rho_\beta}{\partial x_\delta} - \frac{\partial \rho_\delta}{\partial x_\beta} + \frac{\partial(\rho_\beta, W_\delta)}{\partial(v_\gamma, \eta^\gamma)} - \frac{\partial(\rho_\delta, W_\beta)}{\partial(v_\gamma, \eta^\gamma)} \ldots\ldots\ldots\ldots(32).$$

If $O$ describes an infinitesimal closed path, there arises an infinitesimal contact-transformation of the $(\eta, v)$-space, defined by the characteristic functions

$$\Omega = \omega^{\delta\beta}\Omega_{\delta\beta}, \quad R = \rho_{\delta\beta}\,\omega^{\delta\beta} \quad \ldots\ldots\ldots\ldots(33),$$

and we can also write

$$\tfrac{1}{2}\Omega = \xi^\delta \Theta^\beta \Omega_{\delta\beta} = \frac{\partial W(x, \eta, v, \Theta)}{\partial x_\delta}\xi^\delta - \frac{\partial W(x, \eta, v, \xi)}{\partial x_\delta}\Theta^\delta + \frac{\partial W(x, \eta, v, \Theta)}{\partial v_\delta}\frac{\partial W(x, \eta, v, \xi)}{\partial \eta^\delta} - \frac{\partial W(x, \eta, v, \xi)}{\partial v_\delta}\frac{\partial W(x, \eta, v, \Theta)}{\partial \eta^\delta},$$

$$\tfrac{1}{2}R = \frac{\partial \rho(x, \eta, v, \Theta)}{\partial x_\delta}\xi^\delta - \frac{\partial \rho(x, \eta, v, \xi)}{\partial x_\delta}\Theta^\delta + \frac{\partial \rho(x, \eta, v, \Theta)}{\partial v_\delta}\frac{\partial W(x, \eta, v, \xi)}{\partial \eta^\delta} - \frac{\partial \rho(x, \eta, v, \Theta)}{\partial v_\delta}\frac{\partial W(x, \eta, v, \xi)}{\partial \eta^\delta} - \frac{\partial \rho(x, \eta, v, \Theta)}{\partial \eta^\delta}\frac{\partial W(x, \eta, v, \xi)}{\partial v_\delta} + \frac{\partial \rho(x, \eta, v, \Theta)}{\partial \eta^\delta}\frac{\partial W(x, \eta, v, \xi)}{\partial v_\delta} \ldots\ldots\ldots\ldots(34),$$

the $\Omega$ and $R$ being invariant under $P$.

Direct computation gives

$$(\Omega_{\delta\beta})_a + (\Omega_{\beta a})_\delta + (\Omega_{a\delta})_\beta = 0,$$

$$(\rho_{\delta\beta})_a + (\rho_{\beta a})_\delta + (\rho_{a\delta})_\beta = \frac{\partial(\rho_\beta, \Omega_{a\delta})}{\partial(v_\gamma, \eta^\gamma)} + \frac{\partial(\rho_a, \Omega_{\delta\beta})}{\partial(v_\gamma, \eta^\gamma)} + \frac{\partial(\rho_\delta, \Omega_{\beta a})}{\partial(v_\gamma, \eta^\gamma)} \ldots\ldots\ldots\ldots(35).$$

For instance, in Weyl's geometry, where $\rho$ is independent of $\eta$, $v$, the second formula above has zero on the right side. If only $\Omega$ vanishes identically, then, when $O$ describes an infinitesimal closed path, there is no change in the aspect of the subjective world, because the ratios of the single $(\eta, v)$ remain unchanged; if $R$ vanishes, and not $\Omega$, there is only displacement of the figures, but the infinitesimal shape in the $x_a$-space remains unchanged. In the first case, for $\Omega = 0$, the expression

$$\eta^a w_a + v_a \zeta^a \ldots\ldots\ldots\ldots(36)$$

also remains unchanged, and for every pair of $E_{n-1}$-elements of the $O$-bundle, there is one member unchanged by $\Omega$.

10. It is easy to insert the parallel-displacement of Levi-Cività, and the more general point

of view of Weyl in the above theory. But I prefer to specify them as an example of another case. Take an alternating bilinear form as invariant, assume $\rho = 0$, and $n$ even; put

$$a_{\alpha\beta}\eta^\alpha\zeta^\beta = A(\eta, \zeta), \quad a_{\alpha\beta} = -a_{\beta\alpha} \quad\text{.....(37)},$$

the determinant $|a_{\alpha\beta}|$ not being zero. We have for $W$

$$\frac{\partial a_{\alpha\beta}}{\partial x_\gamma}\xi^\gamma\eta^\alpha\zeta^\beta - a_{\alpha\beta}\zeta^\beta\frac{\partial W(x, \eta, v, \xi)}{\partial v_\alpha} - a_{\alpha\beta}\eta^\alpha\frac{\partial W(x, \zeta, w, \xi)}{\partial w_\alpha} = 0 \quad\text{.....(38)};$$

differentiating in regard to $w_\gamma$ we obtain

$$a_{\alpha\beta}\eta^\alpha\frac{\partial^2 W(x, \zeta, w, \xi)}{\partial w_\beta\partial w_\gamma} = 0,$$

so that, the $\eta^\alpha$ being arbitrary, and $|a_{\alpha\beta}|$ not zero, $W$ is linear in the $w_\beta$. If we differentiate with respect to $\zeta^\epsilon$, $\zeta^\delta$, we have, from (38), by the same argument,

$$\frac{\partial^3 W(x, \zeta, w, \xi)}{\partial\zeta^\epsilon\partial\zeta^\delta\partial w_\beta} = 0,$$

shewing that $W$ is also linear in $\zeta$. We therefore put

$$W(x, \eta, v, \xi) = w^\gamma_{\alpha\beta}\eta^\alpha\xi^\beta v_\gamma + u^\gamma_{\alpha\beta}\eta^\alpha\xi^\beta v_\gamma \quad\text{.....(39)},$$

and

$$w^\gamma_{\alpha\beta} = w^\gamma_{\beta\alpha}, \quad u^\gamma_{\alpha\beta} = -u^\gamma_{\beta\alpha};$$

then, comparing coefficients in (38), we get

$$\frac{\partial a_{\alpha\beta}}{\partial x_\gamma} - a_{\delta\beta}w^\delta_{\alpha\gamma} - a_{\delta\beta}u^\delta_{\alpha\gamma} - a_{\alpha\delta}w^\delta_{\beta\gamma} - a_{\alpha\delta}u^\delta_{\beta\gamma} = 0 \quad\text{.....(40)},$$

and, consequently,

$$\frac{\partial a_{\alpha\beta}}{\partial x_\gamma} + \frac{\partial a_{\gamma\beta}}{\partial x_\alpha} + \frac{\partial a_{\alpha\gamma}}{\partial x_\beta} = -(a_{\delta\beta}w^\delta_{\alpha\gamma} + a_{\alpha\delta}w^\delta_{\beta\gamma}) - 2a_{\delta\gamma}u^\delta_{\alpha\beta} \quad\text{.....(41)}.$$

Denoting the subdeterminant of $a_{\delta\gamma}$ in $|a_{\delta\gamma}|$, divided by the determinant itself, by $A^{\delta\gamma}$, we have, multiplying by $A^{\epsilon\gamma}$ and summing,

$$u^\epsilon_{\alpha\beta} = -\tfrac{1}{2}A^{\epsilon\gamma}\left(\frac{\partial a_{\alpha\beta}}{\partial x_\gamma} + \frac{\partial a_{\gamma\beta}}{\partial x_\alpha} + \frac{\partial a_{\alpha\gamma}}{\partial x_\beta}\right) - (a_{\delta\beta}A^{\epsilon\gamma}w^\delta_{\alpha\gamma} + a_{\alpha\delta}A^{\epsilon\gamma}w^\delta_{\beta\gamma}) \quad\text{.....(42)}.$$

Substitution in (40) shews that there is no other condition.

The analogy with the well-known Christoffel symbols is obvious. The symmetrical part of $W$ can be taken arbitrarily, and the alternating part is then defined.

For a symmetrical bilinear form as invariant the results so far mentioned are reciprocal; the symmetrical part of $W$ is determined by the arbitrary alternating part. But $U$ is always an extended point-transformation.

We may also remark that the alternating part of $W$ is invariant under $P$, which modifies only the symmetrical part.

11. Nothing so far given furnishes a measurement in the $x_\alpha$-space. We can obtain such a measurement by considering the one-dimensional strips in which the consecutive $(\eta, v)$-elements arise by a displacement along the $\xi^\alpha$ themselves. The differential equations are

$$\frac{d\eta^\alpha}{dt} = -\frac{\partial W_\beta(x, \eta, v)}{\partial v_\alpha}\eta^\beta + \rho_\beta(x, \eta, v)\eta^\beta\eta^\alpha,$$

$$\frac{dv_\alpha}{dt} = \frac{\partial W_\beta(x, \eta, v)}{\partial\eta^\alpha}\eta^\beta - \rho_\beta(x, \eta, v)\eta^\beta v_\alpha, \quad \frac{dx_\alpha}{dt} = \eta^\alpha \quad\text{.....(43)}.$$

These equations are invariant under $P$, and the parameter $t$ can only be changed into $at+b$, where $a$ and $b$ are constants. If we write down the homogeneous linear partial differential equation corresponding to (43), omitting $t$,

$$\frac{\partial \Phi}{\partial x_a}\eta^a+\frac{\partial \Phi}{\partial \eta^a}\left[-\frac{\partial W_\beta(x,\eta,v)}{\partial v_a}\eta^\beta+\rho_\beta(x,\eta,v)\eta^\beta\eta^a\right]$$

$$+\frac{\partial \Phi}{\partial v_a}\left[\frac{\partial W_\beta(x,\eta,v)}{\partial \eta^a}\eta^\beta-\rho_\beta(x,\eta,v)\eta^\beta v_a\right]=0 \quad\ldots\ldots(44),$$

we see that every integral is an invariant under $W$. For the shapes of the curved lines, which we shall call strips, described by the $x_a$, simultaneously with the $(\eta^a, v_a)$, only those integrals are important which depend on the ratios of the $\eta^a$, and the ratios of the $v_a$; thus we must have

$$\eta^a\frac{\partial \Phi}{\partial \eta^a}=0, \quad v_a\frac{\partial \Phi}{\partial v_a}=0 \quad\ldots\ldots(45);$$

with (44) these two equations form a complete system with $3n-3$ integrals, and, therefore, $3n-3$ arbitrary constants. This number is reduced to $3n-4$ by the condition $\eta^a v_a=0$, which is also an integral of (43). The equations (44), (45) shew that the shape of the strips is independent of the $\rho_\beta$. The definition of the strip from a point $O$ to a point $O'$ requires $2n$ conditions for the $3n-4$ parameters of a strip and the two values of $t$ and $t'$ at $O$ and $O'$; there remain then $n-2$ arbitrary parameters, and we can reach $O'$ from $O$ by moving in a $(n-2)$-fold manifoldness of strips.

If we take an invariant under $W$, say $T(x,\eta,v)$, which is an integral of (44) on account of (13) above, we can fix $t$. And, if $m$ is not zero, we can suppose $m=1$, by taking $T^{\frac{1}{m}}$ instead of $T$. Then, putting

$$T\left(x, \frac{dx_a}{dt}, v_a\right)=1,$$

the parameter $t$ is fixed, and for any strip from $O$ to $O'$ we may put

$$t'-t=\int_O^{O'} T(x, dx_a, v_a),$$

the $x_a$ and $v_a$ being functions of a parameter belonging to the strip.

When $x_a$, $v_a$ pass from $O$ to $O'$ along such a particular strip, defined by (43), we have $\infty^{n-2}$ such strips; we can then choose an extreme one, at least a stationary one, such that the differential $d(t'-t)$, regarded as a function of the $n-2$ parameters, is zero. And this $t'-t$ may be used as a measure of distance between $O$ and $O'$.

Whether such a measurement is additive along the same particular strip, I have not yet ascertained; this would be a very interesting peculiarity. But it is not essential to our purpose; we can, for every strip, determine the particular measurement between two elements which are near enough, and obtain an additive measurement by integration.

12. So many problems arise out of the ideas here explained, that an exhaustive treatment is impossible. We consider now only the application to the Riemann geometry built upon the assumption of a $ds^2$, and the parallel-displacement of Levi-Civitá. For this case the contact-transformation is only an extended point-transformation in the homogeneous space of the $(\eta^a, v_a)$; the particular lines defined by (43) are the ordinary geodesic lines, that joining two points, $O$ and

$O'$, which are sufficiently near, being unique. The $(n-2)$-fold manifoldness at $O$, belonging to $O'$, consists only of the $(n-2)$-fold manifoldness of $E_{n-1}$ containing the initial direction of the geodesic line; every $E_{n-1}$ is carried by parallel-displacement along the geodesic line to $O'$, generating a strip, and reaching $O'$ in the direction of the geodesic line at $O'$. The result is similar for every contact-transformation which is only an extended point-transformation. But if, for $W$, we take a general contact-transformation, we obtain, in the subjective space of an observer at $O$, an $(n-2)$-fold manifoldness, a surface when $n=4$, containing the particular strips coming from $O'$. In the subjective $(n-1)$-dimensional space we have, therefore, an $n$-fold manifoldness of $(n-2)$-dimensional hypersurfaces, forming a representation of the $n$-fold objective manifoldness of points $O'$. If we now imagine some signals, *e.g.* rays of polarised light, coming from the points $O'$, along the particular strips, to the observer at $O$, and we know $W$, we are able to form an image, in part, of the objective world, by observing the $n$-fold manifoldness of $(n-2)$-dimensional hypersurfaces spoken of. For the Riemann geometry, however, and in general for any case of merely extended point-transformation, the observer can only obtain knowledge of events along the geodesic line, and is unable to fix the single point on the line.

13. Can a useful theory of physics be built on these general ideas? It is impossible to affirm or deny. But I cannot finish this paper, already long, without remarking my belief that it should be possible to reconcile the oft-quoted Kantian view of the intuitive character of space and time, with the modern point of view. The observer at $O$ has really a projective three-dimensional space, of the $(\eta^a, v_a)$, and a parameter for his path, in his mind (and that we may call *a priori*); otherwise he may learn something of the objective world by observing the surfaces we have spoken of. But it seems desirable to consider a mathematical scheme as general as possible, both for Geometry and Physics, in order to leave full freedom in the interpretation of observed facts.

# XXIV. *On the Fifth Book of Euclid's Elements.* (*Fifth Paper**.)

By Professor M. J. M. Hill.

[*Received* 14 July. *Read* 31 October 1921.]

1. The object of this paper is to endeavour to recover the train of thought which led the writer (supposed to be Eudoxus, but whom I will refer to as Euclid), of the above-mentioned work, to the formulation of his Fifth Definition (the test for the sameness of two ratios), and his Seventh Definition (the test for distinguishing the greater from the smaller of two unequal ratios).

2. Two attempts to effect the same object are due to De Morgan. The first of these will be found on pp. 25—29 of his treatise on *The Connexion of Number and Magnitude*, to which there is a sub-title *An attempt to explain the Fifth Book of Euclid* (1836). De Morgan considers the distribution of the multiples of a magnitude $A$ amongst the multiples of a magnitude $B$ of the same kind as $A$, thus forming what he calls the relative multiple scale of $A$ and $B$. This he compares with the relative multiple scale of two other magnitudes $C$ and $D$. The argument is intricate requiring the discussion of 81 alternatives.

There is also an attempt, depending on the use of relative multiple scales, to reconstruct the argument by the writer of this paper in the first part of this series.

These arguments are both of a logical nature. On account of the greater simplicity of the procedure which I shall explain below, it seems to me unlikely that Euclid followed either of these methods.

De Morgan's second attempt will be found in the *Penny Cyclopaedia*, Vol. XIX. (1841). A full account is given in Sir T. L. Heath's edition of Euclid's Elements, Vol. II. pp. 122—123. The following is sufficient for comparison with what I now propose.

De Morgan proceeds from *the idea of similarity* and gives the following as an illustration.

Suppose there is a straight colonnade composed of equidistant columns (which may be understood to mean the vertical lines forming the axes of the columns) the first of which is at a distance from a bounding-wall (perpendicular to the straight colonnade) equal to the distance between consecutive columns. In front of the colonnade and parallel to it let there be a straight row of equidistant railings (regarded as meaning their axes), the first being at a distance from the bounding wall equal to the distance between consecutive railings. Let the columns be numbered from the wall and also the railings. The column distance, say $C$, and the railing distance, say $R$, may have any ratio to one another (except that of equality).

Now let a model of the preceding construction be made in which the column distance is $C'$ and the railing distance is $R'$. It needs no definition of proportion, nor anything more than the conception we have of that term prior to any definition (and with which we must show the agreement of any definition that we adopt) to assure ourselves that if the model is truly formed,

* The preceding papers will be found in the 16th, 19th, and 22nd volumes of the *Cambridge Philosophical Transactions*.

then $C$ must have the same ratio to $R$ as $C'$ has to $R'$. Nor is it drawing too largely on that conception of proportion to assert that the distribution of the railings amongst the columns in the model must everywhere be the same as in the original, *i.e.* if the $m$th column be opposite to the space between the $n$th and $(n+1)$th railings in the original, the same must be the case in the model, *i.e.* if $nR < mC < (n+1)R$, then must $nR' < mC' < (n+1)R'$.

If, however, the $s$th column be exactly opposite the $t$th railing in the original, then the same must be the case in the model, *i.e.* if $tR = sC$, then must $tR' = sC'$.

This is equivalent to, though not exactly in the same form as, Euclid's Fifth Definition for the validity of the proportion

$$C : R :: C' : R'.$$

This way of reaching the conditions of the Fifth Definition is simple and direct. It is based on *the idea of similarity.* That idea is of so fundamental a character that several mathematicians from the time of Wallis have proposed that Euclid's Postulate of Parallels should be replaced by the assumption that it is possible to construct a triangle of any size similar to a given triangle. So that it may very well be that Euclid did actually follow this path.

As, however, this method of treating the subject does not afford a simple explanation of the treatment of unequal ratios I venture to make the suggestion that he worked from *the idea of relative magnitude* as his starting-point.

3. In the Third Definition of the Fifth Book ratio is defined as follows (I quote from Sir T. L. Heath's translation):

"A ratio is a sort of relation in respect of size between two magnitudes of the same kind."

There has been a great deal of controversy as to what Euclid meant to imply by this definition, but it does not greatly matter because he makes no use of it in his subsequent argument. The words "of the same kind" are however important. They are used in a technical sense. They mean that two such magnitudes can be added together and the result is a magnitude "of the same kind": that if two such magnitudes are unequal the smaller can be subtracted from the larger and the result is a magnitude "of the same kind"; that if unequal the smaller, if added to itself a sufficient number of times, will give a magnitude "of the same kind" greater than the larger (this is the so-called Axiom of Archimedes); and that any magnitude can be divided into any number of equal magnitudes "of the same kind."

De Morgan, in his Article on Ratio in the *Penny Cyclopaedia, l.c.* p. 308, 1st column, 2nd paragraph, says that ratio is "relative magnitude." It is this *idea of relative magnitude* which has to be explained, and its implications explored.

Euclid nowhere commits himself to the statement that a ratio is itself a magnitude. In his seventh definition he uses the word "greater" in two different senses. The first time it is used it means that one ratio is said to be greater than another provided that a certain condition is satisfied. The second time it is used it is applied to magnitudes in the ordinary sense. Consequently it is held that the first use of the word "greater," when it is applied to ratios, must not be understood in the same way as when it is applied to magnitudes: *e.g.* De Morgan says in his note on this definition that "proof should be given that the same pair of magnitudes can never offer both tests [*i.e.* the test in the definition for a greater ratio and the corresponding test for a less ratio with 'less' substituted for 'greater' in the definition] to another pair; *i.e.* the test of greater ratio from one set of multiples and that of less ratio from another." This is very easily proved, see Heath, *l.c.* p. 130.

4. I suggest that in fact the original exploration of *the idea of relative magnitude* took place on the following lines:

(i) Suppose the magnitudes $A$ and $B$ have a common measure $G$, and that $A = aG$, $B = bG$, where $a$, $b$ are some two whole numbers. If $A$ and $B$ be compared with one another from the point of view of their magnitudes, the first idea that would arise would be that their relation to one another would be the same as that of the whole numbers $a$ and $b$; and that *neither the nature nor the magnitude of their common measure $G$ was material.* If $G$ be not the greatest common measure of $A$ and $B$ then it is always possible to substitute for $G$ that greatest common measure. So in what follows it will be supposed that $G$ is the greatest common measure of $A$ and $B$. Then the magnitudes $A$ and $B$ determine uniquely the whole numbers $a$ and $b$. It would then be laid down AS A DEFINITION that the relative magnitude of $A$ to $B$ was the same as that of the whole number $a$ to the whole number $b$. This is therefore a *definition, not a proposition* as Euclid makes it in X. 5, which reads as follows:

"Commensurable magnitudes have to one another the ratio of a (whole) number to a (whole) number."

Euclid's proof of this proposition depends on the 20th Definition of the Seventh Book and the 22nd Proposition of the Fifth Book.

The result may be expressed thus:

If $$A = aG, \quad B = bG,$$

then the ratio of $A$ to $B$ is the same as that of $a$ to $b$; or in symbols

$$(A : B) = (aG : bG) = (a : b),$$

wherein I have used the symbol for equality in place of Euclid's "is the same as."

Euclid did not take what would now be regarded as the next step, viz.:

The measure of the ratio of $a$ to $b$ is the rational number $a/b$.

(ii) The next step in the train of reasoning was probably the following:

If $$A = aG \text{ and } B = bG,$$

then $$bA = b(aG) = a(bG) = aB,$$

*i.e.* if $A = aG$ and $B = bG$, then $bA = aB$.

(iii) Then would come the attempt to prove the converse proposition.

If $bA = aB$, then $A$ and $B$ must have a common measure.

I think the most likely method adopted for this purpose would be the following:

Assume that $B$ is divided into $b$ equal parts, each equal to $G$.

$$\therefore \; B = bG,$$

$$\therefore \; bA = aB = a(bG) = b(aG),$$

$$\therefore \; A = aG.$$

Hence $A$ and $B$ have a common measure $G$, and therefore $(A : B) = (aG : bG) = (a : b)$.

(For a method of treating this case which does not involve the division of $B$ into equal parts, see Art. 7 (i) below.)

(iv) Then would come the attempt to prove the converse of Euc. x. 5 which is given in Euc. x. 6 "If two magnitudes have to one another the ratio which a (whole) number has to a (whole) number, the magnitudes will be commensurable."

This I think would have been proved as follows:

Given $$(A:B)=(a:b).$$

Divide $B$ into $b$ equal parts, each equal to $G$.

$$\therefore\ B=bG.$$

It would then be argued that the idea of relative magnitude required that

$$(A:B)=(A:bG). \quad \text{[See Section (vi) below.]}$$

But $$(a:b)=(aG:bG) \quad \text{by definition,}$$

$$\therefore\ (A:bG)=(aG:bG).$$

It would then be argued that the idea of relative magnitude required that $A=aG$.

Since $B=bG$ it follows that $A$ and $B$ would have a common measure $G$; and also that $bA=aB$.

We may put this and the preceding conclusion thus:

If $$bA=aB, \text{ then } (A:B)=(a:b).$$

Conversely, if $$(A:B)=(a:b), \text{ then } bA=aB.$$

(v) Having thus reduced the study of the ratios of commensurable magnitudes to the study of the ratios of whole numbers, a closer investigation of these last would be undertaken, and the next set of conclusions that would be reached would be

(α) If $a=b$, then $(a:c)=(b:c)$ and $(c:a)=(c:b)$,

(β) If $a>b$, then $(a:c)>(b:c)$ and $(c:a)<(c:b)$,

which last includes what is set down for symmetry only,

(γ) If $a<b$, then $(a:c)<(b:c)$ and $(c:a)>(c:b)$.

(vi) If now three magnitudes $A$, $B$ and $C$ be taken such that

$$A=aG, \quad B=bG, \quad C=cG,$$

then using the definition in (i) above, it would follow from Section (v) that

if $A=B$, then $(A:C)=(B:C)$ ....................(I),

if $A=B$, then $(C:A)=(C:B)$ ....................(I′),

if $A>B$, then $(A:C)>(B:C)$ ....................(II),

if $A>B$, then $(C:A)<(C:B)$ ....................(II′),

if $A<B$, then $(A:C)<(B:C)$ ....................(III),

if $A<B$, then $(C:A)>(C:B)$ ....................(III′).

In the statement of these six results the fact that $A$, $B$ and $C$ have a common measure, though it is implied, does not obtrude itself.

(vii) It would next be noticed that if (I), (II) and (III) are true, and if we may look upon ratio as a magnitude, then a purely logical deduction leads to the converse propositions,

if $(A:C)=(B:C)$, then $A=B$ ....................(IV),

if $(A:C)>(B:C)$, then $A>B$ ....................(V),

if $(A:C)<(B:C)$, then $A<B$ ....................(VI).

(viii) The next step would be to endeavour to express the ratio of two magnitudes when their common measure was unknown even when it existed. The thinker would be encouraged to make this attempt by the fact noticed under (vi) that the possession of a common measure by $A$, $B$ and $C$ does not obtrude itself in the statement of the conclusions (I), (II), (III), (I′) (II′), (III′).

Suppose that the magnitude $B$ is divided into $b$ equal parts, each equal to $G$, and that it is found that $A$ is intermediate in magnitude between $aG$ and $(a+1)G$.

Then
$$A > aG,$$
$$\therefore\ (A:B) > (aG:B) \quad \text{by (II)},$$
$$(aG:B) = (aG:bG) \quad \text{by (I′)},$$
$$(aG:bG) = (a:b) \quad \text{by definition},$$
$$\therefore\ (A:B) > (a:b).$$

Also
$$A < (a+1)G,$$
$$\therefore\ (A:B) < [(a+1)G:B] \quad \text{by (III)},$$
$$[(a+1)G:B] = [(a+1)G:bG] \quad \text{by (I′)},$$
$$[(a+1)G:bG] = [(a+1):b] \quad \text{by definition},$$
$$\therefore\ (A:B) < [(a+1):b],$$
$$\therefore\ (a:b) < (A:B) < [(a+1):b].$$

Thus the idea would arise that when it was not known whether a common measure of $A$ and $B$ existed, it was nevertheless possible to compare the ratio of $A$ to $B$ with the ratio of one whole number to another.

(ix) The next idea would be that, even when two magnitudes "of the same kind" have no common measure, the idea of relative magnitude involved the truth of the assumptions I, II, III and I′, II′, III′, which had in the first instance been arrived at as a result of the consideration of magnitudes having a common measure.

(x) It would then be surmised that it was possible to determine whether the ratio of one magnitude to another "of the same kind" was greater than, equal to, or less than that of one whole number to another whole number, even when the magnitudes had no common measure.

(xi) It would then be natural to take as the test for the sameness of the two ratios the condition that it was impossible to find the ratio of any whole number to any other whole number which was intermediate in magnitude between the ratios. This would lead in the manner set out in the next article to Euclid's Fifth Definition.

(xii) It would also be natural to take as the test for the inequality of two ratios the condition that it was possible to find the ratio of some whole number to some other whole number which was intermediate in magnitude between the two ratios, or which was equal to one of the ratios but not equal to the other ratio.

5. The whole argument might then be set up as follows:

(i) The ratio of the magnitude $A = aG$ to the magnitude $B = bG$ would be *defined* to be the same as that of $a$ to $b$, and would be written
$$(A:B) = (aG:bG) = (a:b).$$

(ii) It would be assumed as fundamental that the idea of relative magnitude required that, if $A$, $B$, $C$ are magnitudes "of the same kind," whether they have a common measure or not, then

If $A=B$, then $(A:C)=(B:C)$ ......(I),

if $A=B$, then $(C:A)=(C:B)$ ......(I′),

if $A>B$, then $(A:C)>(B:C)$ ......(II),

if $A>B$, then $(C:A)<(C:B)$ ......(II′),

if $A<B$, then $(A:C)<(B:C)$ ......(III),

if $A<B$, then $(C:A)>(C:B)$ ......(III′).

Of these (III) is included in (II) and (III′) in (II′).

(iii) Then it would be deduced as a logical conclusion from I, II and III on the hypothesis that ratios are magnitudes,

if $(A:C)=(B:C)$, then $A=B$ ......(IV),

if $(A:C)>(B:C)$, then $A>B$ ......(V),

if $(A:C)<(B:C)$, then $A<B$ ......(VI).

(iv) The next step would be to compare the ratio of $A$ to $B$ with that of any two whole numbers, say $r$ to $s$.

This might be effected as follows:

Consider the magnitudes $sA$ and $rB$.

It is supposed to be possible to determine whether $sA$ is equal to, or greater than, or less than $rB$.

Suppose that $B$ is divided into $s$ equal parts, each equal to $G$. Then $B=sG$.

Hence $$rB=r(sG)=s(rG).$$

If then
$$sA=rB,$$
$$sA=s(rG),$$
$$A=rG,$$
$$\therefore\ (A:B)=(rG:sG)=(r:s).$$

If however
$$sA>rB,$$
$$sA>s(rG),$$
$$A>rG,$$
$$\therefore\ (A:B)>(rG:B) \quad \text{by (II)},$$
$$(rG:B)=(rG:sG) \quad \text{by (I′)},$$
$$(rG:sG)=(r:s) \quad \text{by definition},$$
$$\therefore\ (A:B)>(r:s).$$

Similarly if $$sA<rB,$$
then $$(A:B)<(r:s).$$

Hence if $sA=rB$, then $(A:B)=(r:s)$ ......(VII),

if $sA>rB$, then $(A:B)>(r:s)$ ......(VIII),

if $sA<rB$, then $(A:B)<(r:s)$ ......(IX).

(v) The propositions converse to VII, VIII and IX might be obtained thus:

Suppose $(A:B)=(r:s)$.

Take as before $B=sG$,

$$\therefore\ (A:B)=(A:sG)\quad \text{by (I')},$$
$$(r:s)=(rG:sG)\quad \text{by definition},$$
$$\therefore\ (A:sG)=(rG:sG),$$
$$\therefore\ A=rG\quad \text{by (IV)},$$
$$\therefore\ sA=s(rG),$$
$$\therefore\ sA=r(sG),$$
$$\therefore\ sA=rB.$$

If however $(A:B)>(r:s)$,

$$(A:B)=(A:sG)\quad \text{by (I')},$$
$$(r:s)=(rG:sG)\quad \text{by definition},$$
$$\therefore\ (A:sG)>(rG:sG),$$
$$\therefore\ A>rG\quad \text{by (V)},$$
$$\therefore\ sA>s(rG),$$
$$\therefore\ sA>r(sG),$$
$$\therefore\ sA>rB.$$

Hence if $(A:B)=(r:s)$, then $sA=rB$ ................................(X).

if $(A:B)>(r:s)$, then $sA>rB$ ................................(XI).

Similarly if $(A:B)<(r:s)$, then $sA<rB$ ................................(XII).

(vi) The next step would be to take as the definition of equal ratios the property that it was impossible to find any two whole numbers such that the ratio of one of them to the other was intermediate in magnitude between the two ratios. Calling the integers $r$ and $s$ and the ratios $(A:B)$ and $(C:D)$,

if $(A:B)=(r:s)$, then must $(C:D)$ be equal to $(r:s)$;

if $(A:B)>(r:s)$, then must $(C:D)$ be greater than $(r:s)$;

if $(A:B)<(r:s)$, then must $(C:D)$ be less than $(r:s)$.

Suppose now $sA=rB$,

then $(A:B)=(r:s)$ by (VII).

Hence we must have $(C:D)=(r:s)$,

$$\therefore\ sC=rD\quad \text{by (X)}.$$

If however $sA>rB$,

then $(A:B)>(r:s)$ by (VIII).

Hence we must have $(C:D)>(r:s)$,

$$\therefore\ sC>rD\quad \text{by (XI)}.$$

Similarly if $sA<rB$,

then $sC<rD$.

So that if $sA=rB$, then must $sC=rD$.

if $sA>rB$, then must $sC>rD$.

if $sA<rB$, then must $sC<rD$.

Now if these conditions be satisfied whatever whole numbers be taken for $r$ and $s$, it will be impossible to find any two whole numbers such that the ratio of one of them to the other is intermediate in magnitude between $(A:B)$ and $(C:D)$, and then these ratios will be said to be equal.

Thus Euclid's Fifth Definition is obtained.

(vii) The next step would be to show how to distinguish between unequal ratios.

Suppose $$(A:B)>(C:D).$$

This could be tested by seeing whether it is possible to find any ratio $(r:s)$ such that

$$(1)\quad (A:B)>(r:s)>(C:D),$$

or $$(2)\quad (A:B)>(r:s)=(C:D),$$

or $$(3)\quad (A:B)=(r:s)>(C:D).$$

These are equivalent respectively to the following:

Integers $r$, $s$ exist such that

$$(1')\quad sA>rB, \quad \text{but } sC<rD,$$

or $$(2')\quad sA>rB, \quad \text{but } sC=rD,$$

or $$(3')\quad sA=rB, \quad \text{but } sC<rD.$$

Euclid's Seventh Definition is equal to the following

$$(A:B)>(C:D)$$

if some integers $r$, $s$ exist such that

$$sA>rB, \text{ but } sC \ngtr rD.$$

This is equivalent to $(1')$ and $(2')$ but Euclid takes no account of $(3')$.

6. I suggest that in the way described above, or in some equivalent way, Euclid reasoned up from the *idea of relative magnitude* to the Fifth and Seventh Definitions; that having obtained them he found that they provided a completely adequate basis for the examination of ratios, and that he then suppressed the argument by which he reached them, possibly in order to avoid the difficulty of treating ratio as a magnitude. He was not in a position to show how to measure ratio because in his time the idea of the irrational number had not been sufficiently developed. It is a matter of controversy as to how far he was in possession of that idea. Whether he actually possessed it or not it is certain that there is the closest connection between his definition of ratios which are the same (or equal) and Dedekind's Theory of Irrational Numbers. In the preface to his tract, entitled "Was sind und was sollen die Zahlen?" Dedekind says that Euclid's Fifth Definition was the source which inspired his theory.

In support of my view that Euclid suppressed the steps by which he reached his 5th and 7th Definitions I would desire to draw attention to the 7th, 8th, 9th and 10th Propositions of the Fifth Book and also the 5th Proposition of the Tenth Book, in which the assumptions numbered I—VI and I′, II′ and III′ and the definition of the ratio of two commensurable magnitudes are proved as propositions based on the 5th and 7th definitions of the Fifth Book and the 20th definition of the Seventh Book. If however any one of the assumptions or the definition of the ratio of two commensurable magnitudes be considered by itself, it seems evident that it is by itself derivable from our idea of relative magnitude, and is of a far more elementary character than the definitions on which its proof in the Fifth or Tenth Book is based.

I would particularly draw attention to the proof in the text of Euclid v. 10, which has been quite rightly criticised by Simson, because Euclid argues in it about greater or less ratios in the same way as if they were magnitudes, though he has not shown that ratios are magnitudes. This seems to me to point to the conclusion that the theory was originally set up in some such way as that described in this paper, but that when the author came to rearrange the argument so as to depend on the 5th and 7th definitions he omitted to cut away as much of the sub-structure as he should have done. Simson attributed the defect in the argument in the 10th (and also that in the 18th proposition) to some later commentator, and explained how these defects could be remedied on Euclid's lines (see Heath, *l.c.* pp. 156—7).

7. It is not without interest to show that the conclusions (VII)—(IX), which have been obtained on the assumption that a magnitude can be divided into any number of equal parts and on the assumptions (I), (II) and (III), can be obtained without assuming the possibility of division into equal parts if we also assume (III′).

This can be effected as follows:

It is convenient to make a slight alteration in the notation.

Let $A$ and $B$ be two magnitudes "of the same kind," and let $a$ and $b$ be any two whole numbers.

Consider the magnitudes $bA$ and $aB$. There are three possibilities.

$$\text{(i)}\ bA = aB \text{ or (ii) } bA > aB \text{ or (iii) } bA < aB.$$

(i) Consider the case* $bA = aB$.

If $a = b$, then $A = B$ and either magnitude is a measure of the other.

If $a \neq b$, suppose $a > b$, then must $A > B$.

Let $A = q_1 B + R_1$, where $q_1$ is a positive integer and $R_1 < B$.

Since $$bA = aB,$$
$$\therefore\ b(q_1 B + R_1) = aB,$$
$$\therefore\ bR_1 = (a - bq_1) B.$$

Put $a - bq_1 = r_1$, $$\therefore\ bR_1 = r_1 B.$$

Now $R_1 < B$, $\therefore\ r_1 < b$.

Let $B = q_2 R_1 + R_2$, where $q_2$ is a positive whole number and $R_2 < R_1$.

$$\therefore\ r_1(q_2 R_1 + R_2) = bR_1.$$
$$\therefore\ r_1 R_2 = (b - q_2 r_1) R_1.$$

Put $b - q_2 r_1 = r_2$. $$\therefore\ r_1 R_2 = r_2 R_1.$$

Now $R_2 < R_1$, $\therefore\ r_2 < r_1$.

Hence $$A = q_1 B + R_1, \quad a = q_1 b + r_1,$$
$$B = q_2 R_1 + R_2, \quad b = q_2 r_1 + r_2.$$

and so on.

It is obvious that this process amounts to finding simultaneously the greatest common measure of $A$ and $B$ and that of $a$ and $b$.

* The examination of this case was given me several years ago by Mr Rose-Innes.

But $a$ and $b$ being positive whole numbers the process must necessarily come to an end.

Suppose that it is found that $r_{n-1} = q_{n+1} r_n$.

Then we have at the same time $R_{n-1} = q_{n+1} R_n$.

If now $a$ and $b$ have a common factor it may be supposed to be divided out from both sides of the equation $bA = aB$, so that $a$ and $b$ may be regarded as prime to one another, and then their greatest common measure, $r_n$, will be unity.

Moreover $A$ will be the same multiple of $R_n$ as $a$ is of $r_n$, i.e. unity.

$$\therefore \; A = aR_n.$$

Similarly $B$ will be the same multiple of $R_n$ as $b$ is of $r_n$, i.e. unity.

$$\therefore \; B = bR_n.$$

$$\therefore \; (A : B) = (aR_n : bR_n) = (a : b).$$

Hence if $bA = aB$, then $(A : B) = (a : b)$.

(ii) Consider next the case $bA > aB$.

Let $bA - aB = C$.

It will first be proved that a magnitude $D$ exists such that $abD < C$.

Suppose that $E$ is any magnitude of the same kind as $A$ and $B$.

Then, as in Euc. X. 1, suppose that the remainder left after taking away from $E$ its half or more than its half is $E_1$.

Let the remainder left after taking away from $E_1$ its half or more than its half be $E_2$. Let this process be repeated $n$ times and let the remainder be $E_n$.

Now consider the magnitudes

$$abE, \; abE_1, \; abE_2, \; \ldots, \; abE_n.$$

Then each magnitude is the remainder left after taking away from the preceding its half or more than its half.

It follows from Euc. X. 1 that, if this process be carried on far enough, there will at length be left a remainder less than any assigned magnitude.

Suppose then that $abE_n < C$. Then the magnitude $E_n$ can be taken to be $D$, and so the existence of $D$ is proved.

Now
$$bA = aB + C,$$
$$C > abD,$$
$$\therefore \; bA > aB + abD.$$

Put for brevity $bA = X$, $aB = Y$, $abD = Z$,

$$\therefore \; X > Y + Z,$$
$$\therefore \; X > Z \text{ and } X - Z > Y.$$

Now form the successive multiples of $Z$, viz.:

$$Z, \; 2Z, \; 3Z, \; \ldots.$$

Suppose $tZ$ the greatest multiple of $Z$ which is less than $X$.

Then either $(t+1)Z = X > tZ$ or $(t+1)Z > X > tZ$,

$$\therefore \; tZ = X - Z > Y \text{ or } tZ > X - Z > Y.$$

In both cases therefore $X > tZ > Y$.

Replacing $X$, $Y$ and $Z$ by their values it follows that

$$bA > tabD > aB.$$

Put now $taD = A'$ and $tbD = B'$,

$$\therefore\ bA > bA' \text{ and } aB' > aB.$$

$$\therefore\ A > A' \text{ and } B' > B.$$

Since $A > A'$ $\therefore\ (A : B) > (A' : B)$ by (II).

Since $B < B'$ $\therefore\ (A' : B) > (A' : B')$ by (III').

$$\therefore\ (A : B) > (A' : B) > (A' : B').$$

But $A' = a(tD)$ and $B' = b(tD)$,

$$\therefore\ (A' : B') = (a : b),$$

$$\therefore\ (A : B) > (a : b).$$

(iii) Similarly it can be shown that

$$\text{if } bA < aB \text{ then } (A : B) < (a : b).$$

Hence if $bA = aB$ then $(A : B) = (a : b)$ which is (VII),

if $bA > aB$ then $(A : B) > (a : b)$ which is (VIII),

and if $bA < aB$ then $(A : B) < (a : b)$ which is (IX).

8. In Article 5 it was pointed out that the condition that one ratio may be greater than another may take one of three different forms. The second and third forms can occur only when one at least of the ratios is that of commensurable magnitudes. It can be shown that in either of these two cases other integers $r'$ and $s'$ can be found so that the condition can be replaced by one of the first form.

Take the second form.

Suppose that $sA > rB$, $sC = rD$.

Since $sA > rB$ it follows from Archimedes' Axiom that an integer $n$ exists such that

$$n(sA - rB) > B,$$

$$\therefore\ nsA > (nr + 1)B,$$

but

$$sC = rD,$$

$$\therefore\ nsC = nrD < (nr + 1)D.$$

Put $ns = s'$, $nr + 1 = r'$,

$$\therefore\ s'A > r'B,\ s'C < r'D.$$

This is of the first form.

The third form can be treated in like manner.

## *Note on the preceding papers.*

This series of papers originated in the discovery that all the propositions of the Fifth Book concerning properties of Equal Ratios could be deduced from the Fifth Definition (the test for the equality of ratios) without using (as Euclid does in some of these propositions) propositions which depend on the Seventh Definition (the test for distinguishing between unequal ratios).

Proofs of the propositions on these lines were given in the first paper of this series and also in the *First* Edition of my Contents of the Fifth and Sixth Books of Euclid (Cambridge University Press, 1900). Most of these proofs are deduced directly from the Fifth Definition *without using any other proposition.*

In the proofs of Euc. v. 9 (Part I), 16, 22 and 23 the Axiom of Archimedes is employed. The proofs given of Euc. v. 9 (Part II), 19, 24 and 25 do however depend on other propositions.

The second part of Euc. v. 9 can however be proved in much the same way as the first part is proved in the first paper of this series, and it is not necessary to make use of the Corollary to Euc. v. 4. Hence Euc. v. 9 can be classed with those propositions which can be deduced directly from the Fifth Definition without using any other proposition.

Somewhat complicated proofs of Euc. v. 19, 24 and 25 depending on the Fifth Definition but not depending on other propositions (with the exception in one case of a lemma) were given in the second paper of this series; and less intricate proofs of the same propositions, due to Mr Rose-Innes, were given in the third paper of this series.

Euc. v. 19 is a transformation of Euc. v. 17 by means of Euc. v. 16, and Euc. v. 25 is a simple application of Euc. v. 19. I believe that no simpler proofs than those in the Fifth Book can be constructed. It seems necessary in these two cases to use other properties of equal ratios if complexity is to be avoided.

There remains Euc. v. 24. The proof given below is much longer than Euclid's proof. On the other hand it proceeds on the same direct lines as the proofs given in my preceding work for Euc. v. 16, 22 and 23. It does not depend on other propositions. The steps follow one another in a natural order; they do not require the amount of search for the next step at each stage in the argument that Euclid must have employed.

I will set the proof out, not in the manner I employed in my first paper, but in that used in the *Second* Edition of my Contents of the Fifth and Sixth Books of Euclid (1908) and in my Theory of Proportion (Constable & Co., 1914), and accordingly I will employ the fractional notation for the measure of the ratio of two whole numbers.

The proposition is to show that

$$\text{if } (A : C) = (X : Z),$$

and

$$\text{if } (B : C) = (Y : Z),$$

then

$$[(A + B) : C] = [(X + Y) : Z].$$

Compare $[(A + B) : C]$ with *any* rational fraction *whatever*, say $s/r$. Then it is known from previous work [Art. 5 of my Third Paper and Art. 48 of the *Second* Edition of my Contents of the Fifth and Sixth Books of Euclid] that it is sufficient to consider only the two alternatives:

(i) $[(A+B):C] > s/r$,

$\therefore\ r(A+B) > sC$.

Choose $n$ so large that

$n[r(A+B) - sC] > 2C$*.

Let $uC$ and $vC$ be the *greatest* multiples of $C$ which are less than $nrA$, $nrB$ respectively; and suppose $n$ so large that $nrA$ and $nrB$ are each greater than $C$.

$\therefore\ nrA - uC \gtreqless C$.

$nrB - vC \gtreqless C$,

$\therefore\ nr(A+B) - (u+v)C \gtreqless 2C$,

but $nr(A+B) - nsC > 2C$,

$\therefore\ (u+v)C > nsC$,

$\therefore\ u+v > ns$.

Now $nrA > uC$,

and $(A:C) = (X:Z)$,

$\therefore\ nrX > uZ$.

Also $nrB > vC$,

and $(B:C) = (Y:Z)$,

$\therefore\ nrY > vZ$,

$\therefore\ nr(X+Y) > (u+v)Z$,

but $u+v > ns$,

$\therefore\ nr(X+Y) > nsZ$,

$\therefore\ r(X+Y) > sZ$,

$\therefore\ [(X+Y):Z] > s/r$.

Hence if

$[(A+B):C] > s/r$,

then $[(X+Y):Z] > s/r$.

or (ii) $[(A+B):C] < s/r$,

$\therefore\ r(A+B) < sC$.

Choose $n$ so large that

$n[sC - r(A+B)] > 2C$*.

Let $uC$ and $vC$ be the *least* multiples of $C$ which are greater than $nrA$, $nrB$ respectively.

$\therefore\ uC - nrA \gtreqless C$.

$vC - nrB \gtreqless C$,

$\therefore\ (u+v)C - nr(A+B) \gtreqless 2C$,

but $nsC - nr(A+B) > 2C$,

$\therefore\ nsC > (u+v)C$,

$\therefore\ ns > u+v$.

Now $nrA < uC$,

and $(A:C) = (X:Z)$,

$\therefore\ nrX < uZ$.

Also $nrB < vC$,

and $(B:C) = (Y:Z)$,

$\therefore\ nrY < vZ$,

$\therefore\ nr(X+Y) < (u+v)Z$,

but $u+v < ns$,

$\therefore\ nr(X+Y) < nsZ$,

$\therefore\ r(X+Y) < sZ$,

$\therefore\ [(X+Y):Z] < s/r$.

Hence if

$[(A+B):C] < s/r$,

then $[(X+Y):Z] < s/r$.

Hence $s/r$ does not lie between the ratios

$$[(A+B):C] \text{ and } [(X+Y):Z].$$

But $s/r$ represents *any* rational fraction *whatever*.

Therefore *no* rational fraction *whatever* lies between these ratios.

$$\therefore\ [(A+B):C] = [(X+Y):Z].$$

Thus this proposition may be included amongst those properties of equal ratios which are directly deducible from the Fifth Definition *without using any other property of equal ratios.*

* The fact that on the right-hand side of these two inequalities there appears $2C$ and not $C$ makes a difference between this proof and those of Euc. v. 16, 22 and 23 and makes it more difficult. It is certainly not obvious at first sight. The reason for its necessity comes out in the proof.

The essential difference between the proofs given in my papers and those given by Euclid of Props. 16, 22 and 23 consists in this:

Instead of using Euclid's Seventh Definition and his Propositions regarding Unequal Ratios, viz.: 8, 10 and 13, I use a proposition regarding multiples of magnitudes, which is included in the earlier part of Euc. v. 8, viz.:

If $A$, $B$, $C$ be three magnitudes "of the same kind" and if $A$ be greater than $B$, then integers $n$ and $t$ exist such that

$$nA > tC > nB.$$

To prove this Euclid uses the Axiom of Archimedes for the first and only time in his Fifth Book.

Since $A > B$,

$\therefore$ $A - B$ is a magnitude of the same kind as $C$,

$\therefore$ by Archimedes Axiom an integer $n$ exists such that

$$n(A - B) > C,$$

$$\therefore\ nA > nB + C.$$

Now put $$nA = X,\quad nB = Y,\quad C = Z,$$

$$\therefore\ X > Y + Z.$$

Then as in Art. 7 (ii) of this paper an integer $t$ exists such that

$$X > tZ > Y,$$

$$\therefore\ nA > tC > nB.$$

In order to prove Euc. v. 16, 22 and 23 it is far simpler to use this proposition than Euc. v. 8, 10 and 13 which depend on it. Props. 14, 20 and 21, which are particular cases of Props. 16, 22 and 23, need not then be proved before Props. 16, 22 and 23 can be obtained.

# XXV. *The influence of electrically conducting material within the earth on various phenomena of terrestrial magnetism.*

By S. Chapman, M.A., Trinity College and T. T. Whitehead.

(with two figures)

[*Received* 16 March, *Read* 1 May, 1922.]

## CONTENTS.



## INTRODUCTION.

§ 1. In his memoir of 1889 on the diurnal magnetic variations Sir A. Schuster* found that these mainly originate above the earth's surface (in the atmosphere, in fact), but that there is also a part proceeding from within the earth. The latter part he regarded as induced by a primary external varying magnetic field, acting upon conducting material within the earth. In this section of the discussion he was assisted by Prof. H. Lamb†, who contributed to the memoir a mathematical appendix on electromagnetic induction in spheres. The assumption which was naturally first considered was that for this purpose the earth might be regarded as a uniformly conducting sphere, but the hypothesis proved irreconcilable with the facts; the amplitude-ratio and phase-difference between the external and internal portions of the observed field did not agree with any possible pair of values deducible from this hypothesis. It was pointed out, however, that agreement could be obtained by supposing that only a concentric core of the earth was conducting, the outer portion taking no share in the phenomenon. The data were inadequate to test the extended hypothesis in detail, and this was first done by one of the present writers, using more observational material‡. Numerical estimates were made of the size and conductivity of the core, assuming that its magnetic permeability is unity. The thickness of the non-conducting shell surrounding the core was estimated at about 250 kilometres, while the value obtained for the conductivity of the core was $3{\cdot}6 \, . \, 10^{-13}$ C.G.S. units. This is of the same order as that of moist earth, and distinctly less than that of sea-water $(4 \, . \, 10^{-11})$.

* A. Schuster, *Phil. Trans.* A 180, p. 467, 1889.

† H. Lamb, Appendix to the above.

‡ S. Chapman, *Phil. Trans.* A 218, p. 1, 1919.

This model of the earth can only be regarded as a convenient first approximation, particularly since the oceans and the water-bearing land strata near the earth's surface are ignored. It seemed desirable to examine to what extent these surface conducting layers would influence the induced magnetic field. This necessitated an extension of Prof. Lamb's analysis, in order to determine the relation between the magnetic field on the two sides of a spherical shell, when on both sides the field arises partly beneath (within) and partly above (outside) the region considered. This extension, which is quite straightforward, is made in Part I, the results being discussed in Part II. No attempt is made to deal with the actual distribution of land and ocean, the surface conducting shell being supposed uniform and complete, separated from a uniform core by a non-conducting shell. It is found that the influence of any probable depth of moist earth is almost negligible, but that a comparatively shallow oceanic shell produces induction effects comparable with those of the supposed core. It is suggested that on this account there should be observable differences between the diurnal magnetic variation at continental and oceanic stations, though in view of the irregular distribution of land and water the calculation of the differences would be arduous. A rough attempt is made, however, to estimate the change in the conductivity and radius of the core when use is made only of continental magnetic data (§ 11).

Some related problems are also considered, in particular, the magnitude and type of the earth currents flowing near the earth's surface, both those which accompany the ordinary diurnal magnetic variations, and the larger currents observed when the earth's field is varying rapidly and irregularly. Again, the main symmetrical part of the field of a magnetic storm is examined, and it is shown that the relation between the horizontal and vertical components is compatible with the existence of a core, the conductivity of which is of the order inferred from the study of the diurnal variation. Finally, the opportunity is taken of seeing how far the original estimate of the conductivity ($\kappa$) and the size of the core would be modified if the unlikely assumption were made that the permeability ($\mu$) differs appreciably from unity. It appears that, consistently with the observed data, a wide range of values is permissible for $\kappa$ provided that $\mu$ varies almost proportionately (and this without much change in the thickness of the outer non-conducting shell). It is only natural that $\kappa$ should have to be larger if $\mu$ is larger, for the inducing field penetrates less far into the core, and the induced currents have to flow in a layer of diminished depth.

## PART I

*Mathematical theory of induction in a spherically symmetrical earth.*

§ 2. It is convenient to tabulate the following symbols for convenience of reference.

VECTORS

**A** = magnetic vector potential

**B** = magnetic induction

**H** = magnetic force

**E** = electric force.

Spherical polar components of these vectors, reckoned positive in the direction of increasing values of $r$, $\theta$, $\phi$, will be denoted by the corresponding letter as suffix. The positive directions for $\theta$ and $\phi$ are from the north pole, and eastwards, respectively.

SCALARS

$\mu$ = permeability

$\kappa$ = conductivity

$c$ = electromagnetic constant $3 . 10^{10}$

$E, I$ = complex constants representing in amplitude and phase the intensity of the parts of a magnetic field originating respectively outside and beneath the region occupied by the point to which the intensities $E$ and $I$ refer. Suffixes $s$, $o$, and $i$ are added to these letters, to indicate which is the region concerned: $s$ refers to a point in the substance of a conducting sphere or shell, $o$ to a region outside (above) such a conducting body, and $i$ to the space within (below) the inner surface of a conducting shell.

Except where the contrary is stated, the units used are electrostatic.

§ 3. The vector potential **A** is defined by the equations

$$\mathbf{B} = \text{curl } \mathbf{A}, \qquad \text{div } \mathbf{A} = 0. \qquad \ldots\ldots 3{\cdot}1$$

If $N$ is the number of tubes of induction through any circuit $s$ enclosing an area $f$,

$$N = \int_f B_n \, df = \int_s A_s \, ds = \int_s \mathbf{A} \, \mathbf{ds}$$

by Stokes' theorem.

But

$$\int_s \mathbf{E} \, \mathbf{ds} = -\frac{1}{c}\frac{dN}{dt} = -\frac{1}{c}\frac{d}{dt}\int_s \mathbf{A} \, \mathbf{ds}$$

if electrostatic effects are neglected, so that

$$\mathbf{E} = -\frac{1}{c}\frac{d\mathbf{A}}{dt}.$$

Also

$$\text{curl } \mathbf{B} = \text{curl curl } \mathbf{A} = -\nabla^2 \mathbf{A} + \text{grad div } \mathbf{A}$$
$$= -\nabla^2 \mathbf{A}$$

since div **A** is zero. Again, since

$$\text{curl } \mathbf{B} = \mu \text{ curl } \mathbf{H} = \frac{4\pi\mu\kappa}{c} \mathbf{E} = -\frac{4\pi\mu\kappa}{c^2} \cdot \frac{d\mathbf{A}}{dt}$$

it follows that

$$\nabla^2 \mathbf{A} = \frac{4\pi\mu\kappa}{c^2} \cdot \frac{d\mathbf{A}}{dt}. \qquad \ldots\ldots 3{\cdot}2$$

In non-conducting material this becomes

$$\nabla^2 \mathbf{A} = 0. \qquad \ldots\ldots 3{\cdot}3$$

§ 4. An appropriate solution of equation 3·2 is sought. A field of external origin with a given magnetic potential near the earth's surface will be considered, this potential being supposed analysed into spherical harmonic components about the geographical axis of the earth. The typical term of this potential will be of the form

$$\Omega_n \equiv E^p{}_{n,a} \frac{r^n}{a^{n-1}} P_n{}^p (\cos\theta) \, e^{\iota p\phi} e^{\alpha t}, \qquad \ldots\ldots 4{\cdot}1$$

where $E$ is the complex constant already mentioned,

$a$ is the earth's radius,

$P_n{}^p$ is the usual associated Legendre function, and

$t$ is the time of some standard meridian.

As regards $\alpha$, three principal special cases will be considered, viz., those of (i) the diurnal magnetic variations, supposed dependent solely on local time, so that $\alpha = \iota p$, enabling* $t+\phi$ to be replaced by the local time $t'$; (ii) fairly rapid and local periodic variations, in which $\alpha$ is a pure imaginary $\iota m$ distinct from $\iota p$ (the local nature of the variations is indicated by giving $m$ and $p$ fairly large values, say 8 to 10 or more—cf. § 13); and (iii) non-periodic variations, such as those of the main worldwide component of a magnetic storm, which can be represented by a combination of terms in which $p=0$, and $\alpha$ is real (§ 12).

If this field of external origin induces a current system within the earth, the typical term of the potential of the secondary field will be

$$\Omega_{-n-1} \equiv I^p_{n,\alpha} \frac{a^{n+2}}{r^{n+1}} P_n^p (\cos\theta)\, e^{\iota p\phi}\, e^{\alpha t}. \qquad \text{4·2}$$

§ 5. It may be readily verified that in non-conducting space the terms in the vector-potential corresponding to $\Omega_n$ and $\Omega_{-n-1}$, are as follows:

$$A_r = 0 \qquad A_r = 0$$

$$A_\theta = -\frac{1}{n+1}\frac{\partial\Omega_n}{\partial\lambda} \qquad A_\theta = \frac{1}{n}\frac{\partial\Omega_{-n-1}}{\partial\lambda}$$

$$A_\phi = \frac{1}{n+1}\frac{\partial\Omega_n}{\partial\theta} \qquad A_\phi = -\frac{1}{n}\frac{\partial\Omega_{-n-1}}{\partial\theta}, \qquad \text{5·10, 5·11}$$

where

$$\partial\lambda = \sin\theta\,\partial\phi. \qquad \text{5·2}$$

These values of **A** satisfy 3·3 (since $\nabla^2\Omega_n$ and $\nabla^2\Omega_{-n-1}$ are both zero) and the components of curl **A** equal the corresponding components of **B**.

In a region occupied by conducting matter, the appropriate solution for **A** is found to be

$$A_r = 0 \qquad A_r = 0$$

$$A_\theta = -\frac{\partial u_n}{\partial\lambda} \qquad A_\theta = -\frac{\partial u_{-n-1}}{\partial\lambda}$$

$$A_\phi = \frac{\partial u_n}{\partial\theta} \qquad A_\phi = \frac{\partial u_{-n-1}}{\partial\theta} \qquad \text{5·30, 5·31}$$

for the terms of external and internal origin respectively, provided that in either case $u$ satisfies the equation†

$$\nabla^2 u = \frac{4\pi\kappa\mu}{c^2}\cdot\frac{du}{dt}. \qquad \text{5·4}$$

The terms in $u$ corresponding to the typical terms 4·1, 4·2 in the magnetic potential are both of the form

$$u = f(r)\, P_n^p\, e^{\iota p\phi}\, e^{\alpha t}, \qquad \text{5·5}$$

where $f(r)$, a function of $r$ only, is a solution of the equation

$$\frac{\partial}{\partial r}\left\{r^2\frac{\partial}{\partial r}f(r)\right\} - \{n(n+1)+k^2r^2\} f(r) = 0, \qquad \text{5·6}$$

the new constant $k$, introduced for brevity, being defined by

$$k^2 = \frac{4\pi\kappa\mu\alpha}{c^2}. \qquad \text{5·7}$$

* Provided the unit in which $t$ is measured is $86400/2\pi$ or $89500/2\pi$ seconds for the solar and lunar diurnal variations respectively; $p$ is then the number of periods per day.

† The solution 5·3 agrees with the solution 5·1 when $\kappa=0$, the actual values of $u$ being $\Omega_n/(n+1)$ and $-\Omega_{-n-1}/n$.

The components of **B** or curl **A** corresponding to 5·5 are

$$B_r = -n(n+1)u/r, \quad B_\theta = -\frac{1}{r}\cdot\frac{\partial}{\partial\theta}\cdot\frac{\partial(ru)}{\partial r}, \quad B_\phi = -\frac{1}{r\sin\theta}\cdot\frac{\partial}{\partial\phi}\cdot\frac{\partial(ru)}{\partial r}, \quad \text{........5·8}$$

whether the origin be external or internal.

§ 6. The equation 5·6 is analytic at all finite points except $r=0$, which is a regular point. The indicial equation has the solutions $n$, $-n-1$, so that $f(r)$ must be of the form

$$Ar^n R_n + Br^{-n-1} R_{-n-1}, \quad \text{..........6·1}$$

where $A$, $B$ are arbitrary constants, and $R_n$, $R_{-n-1}$ are integral functions satisfying the equations

$$r^2\frac{\partial^2 R_m}{\partial r^2} + 2(m+1)r\frac{\partial R_m}{\partial r} - k^2r^2R_m = 0. \quad \begin{Bmatrix} m=n \text{ or} \\ m=-n-1 \end{Bmatrix}. \quad \text{......6·2}$$

Clearly $R_n$ and $R_{-n-1}$ are functions of $kr$ only; in the following sections this value of the argument is intended to be understood when no other is specially indicated.

If the field is of purely external origin, $B=0$; while if it is of purely internal origin, $A=0$.

By actual substitution it is found that

$$R_n = 1 + \frac{k^2r^2}{2(2n+3)} + \frac{k^4r^4}{2.4(2n+3)(2n+5)} + \dots, \quad \text{..........6·3}$$

whence the following useful recurrence formulae can be readily obtained:

$$r\frac{\partial R_n}{\partial r} = \frac{k^2r^2}{2n+3}R_{n+1}, \quad \text{..........6·4}$$

$$R_{n+1} = \frac{(2n+1)(2n+3)}{k^2r^2}(R_{n-1} - R_n). \quad \text{..........6·5}$$

When $n=0$ the series 6·3 reduces to a specially simple form, giving

$$R_0 = \frac{\sinh kr}{kr}, \quad \text{..........6·6}$$

so that, by 6·4,

$$R_1 = \frac{3}{k^2r^2}\left(\cosh kr - \frac{\sinh kr}{kr}\right), \quad \text{..........6·7}$$

while by further applications of 6·4 the expression for $R_n$, whatever the value of $n$, can be obtained as the sum of a finite number of hyperbolic functions, equivalent to the series 6·3.

The function $R_{-n-1}$ can be expressed most simply in the form $e^{-kr}\sigma(r)$, where $\sigma$ is the integral solution of the equation

$$r^2\frac{\partial^2\sigma}{\partial r^2} - 2r(n+kr)\frac{\partial\sigma}{\partial r} + 2nkr\sigma = 0. \quad \text{..........6·8}$$

It is found on substitution that the series for $\sigma(r)$ terminates with the power $r^n$, the actual result being

$$R_{-n-1} = e^{-kr}\left\{1 + kr + \frac{2(n-1)k^2r^2}{2(2n-1)} + \frac{2^2(n-2)(n-1)k^3r^3}{3!(2n-2)(2n-1)} + \dots\right\}, \quad \text{..........6·9}$$

unless $n=0$, when $$R_{-1} = e^{-kr}. \quad \text{..........6·10}$$

Clearly $$R_{-2} = e^{-kr}(1+kr). \quad \text{..........6·11}$$

From 6·9 the following formulae, corresponding to 6·4 and 6·5, can be deduced:

$$r\frac{\partial R_{-n-1}}{\partial r}=\frac{k^2r^2}{1-2n}R_{-n}, \qquad \text{6·12}$$

$$R_{-n-1}=R_{-n}+\frac{k^2r^2}{(2n-1)(2n-3)}R_{-n+1}. \qquad \text{6·13}$$

When the real part of $kr$ is sufficiently large, $\sinh kr$ and $\cosh kr$ can be replaced by $\frac{1}{2}e^{kr}$; repeated use of 6·4 then gives the approximate formula

$$R_n=\tfrac{1}{2}e^{kr}\frac{(2n+1)!}{2^n.n!(kr)^{n+1}}\left\{1-\frac{n(n+1)}{2kr}+\frac{(n-1)n(n+1)(n+2)}{2^2.2!(kr)^2}-\frac{(n-2)\dots(n+3)}{2^3.3!(kr)^3}+\dots\right\} \qquad \text{6·14}$$

The corresponding formula for $R_{-n-1}$ follows from 6·9 merely by algebraic transformation, viz.,

$$R_{-n-1}=\tfrac{1}{2}e^{-kr}\frac{2^n(n-1)!(kr)^n}{(2n-1)!}\left\{1+\frac{n(n+1)}{2kr}+\frac{(n-1)n(n+1)(n+2)}{2^2.2!(kr)^2}+\dots\right\}. \qquad \text{6·15}$$

It may be noted that the last two formulae are really asymptotic expansions of Bessel's functions*, multiplied by a factor.

It will be necessary later (§ 7) to calculate numerical values of $R_{n-1}/R_n$, in cases where $kr$ is large; it may be deduced from 6·14 that

$$\frac{R_{n-1}}{R_n}=\frac{kr}{2n+1}\left\{1+\frac{n}{kr}+\frac{n(n+1)}{2(k^2r^2)}+\frac{n(n+1)}{2(k^3r^3)}+\dots\right\}. \qquad \text{6·16}$$

§ 7. *Induction in a uniform conducting permeable sphere.*

The above analysis will first be applied to the case of a uniform spherical core of magnetic permeability $\mu_s$ and conductivity $\kappa$, in which an external periodic magnetic field depending only on local time (so that $\alpha=\iota p$; cf. § 4) induces a secondary internal field. The radius of the core will be taken as $qa$ ($q \leqslant 1$), and the ratio of the corresponding terms in the potentials of the two fields will be determined, for a point at the earth's surface ($r=a$) in non-conducting material surrounding the core. The required ratio is $\Omega_{-n-1}/\Omega_n$ for $r=a$, and by 4·1, 4·2 this equals $I_o/E_o$.

Outside the core

$$H_r=\{-nE_o(r/a)^{n-1}+(n+1)I_o(a/r)^{n+2}\}P_n^pe^{\iota pt'}, \qquad \text{7·1}$$

$$H_\phi=-\iota p\,\mathrm{cosec}\,\theta\{E_o(r/a)^{n-1}+I_o(a/r)^{n+2}\}P_n^pe^{\iota pt'} \qquad \text{7·2}$$

and similarly for $H_\theta$; $t'(\equiv t+\phi)$ denotes the local time in angular measure.

The corresponding values of the magnetic induction within the core are

$$B_r=-n(n+1)E_s(r/a)^{n-1}R_nP_n^pe^{\iota pt'}, \qquad \text{7·3}$$

$$B_\phi=-\iota pE_s(r/a)^{n-1}\mathrm{cosec}\,\theta\left\{(n+1)R_n+r\frac{\partial R_n}{\partial r}\right\}P_n^pe^{\iota pt'} \qquad \text{7·4}$$

(and similarly for $B_\theta$), since there is no primary field within the core.

At the surface of the core ($r=qa$) the normal induction and tangential force are continuous, so that

$$nq^{n-1}E_o-(n+1)q^{-n-2}I_o=n(n+1)q^{n-1}E_sR_n, \qquad \text{7·5}$$

$$\mu_s(q^{n-1}E_o+q^{-n-2}I_o)=q^{n-1}E_s\left\{(n+1)R_n+qa\frac{\partial R_n}{\partial r}\right\}. \qquad \text{7·6}$$

* Cf. G. N. Watson, *Proc. Roy. Soc.* A 95, p. 83, 1918.

Hence, using 6·4, and eliminating $E_s$, it appears that

$$\frac{I_o}{E_o}=\frac{n}{n+1}q^{2n+1}\left\{\frac{k^2q^2a^2R_{n+1}+(2n+3)(n+1)(1-\mu_s)R_n}{k^2q^2a^2R_{n+1}+(2n+3)(n+\mu_s n+\mu_s)R_n}\right\}. \quad\text{......7·7}$$

In the last three equations, and in 7·8, 7·9, the value of $r$ to be substituted in $R_n$ or $R_{n+1}$ is $qa$.

Equation 7·7 reduces to the result previously obtained by Prof. Lamb on putting $q=\mu_s=1$.

By means of 6·5 equation 7·7 can be written in the form

$$\frac{I_o}{E_o}=\frac{n}{n+1}q^{2n+1}\left[1-\mu_s\left\{\frac{R_{n-1}}{R_n}+\frac{n(\mu_s-1)}{2n+1}\right\}^{-1}\right], \quad\text{......7·8}$$

which reduces to

$$\frac{I_o}{E_o}=\frac{n}{n+1}q^{2n+1}\left(1-\frac{R_n}{R_{n-1}}\right), \quad\text{......7·9}$$

if $\mu_s=1$.

In the present case $k^2r^2$ is a purely imaginary quantity (cf. 5·7, in which $\iota p$ has now to be substituted for $\alpha$); it is convenient to write

$$k^2q^2a^2=2i\beta^2 \quad\text{......7·10}$$

so that

$$kr=\beta(1+i),\qquad \frac{1}{kr}=\frac{1-i}{2\beta}. \quad\text{......7·11}$$

On substitution of these values the equation 6·16 becomes

$$\frac{R_{n-1}}{R_n}=\frac{\beta}{2n+1}\left[\left\{1+\frac{n}{\beta}+\frac{n(n+1)}{4\beta^2}+\frac{n(n+1)}{4\beta^3}+\ldots\right\}+i\left\{1-\frac{n(n+1)}{4\beta^2}+\ldots\right\}\right], \quad\text{......7·12}$$

where $n$ and $\beta$ are real and positive. The other formulae of § 6 can likewise be readily obtained in terms of $\beta$, though actual values of $R_n$ and $R_{-n-1}$ are not required in this paper.

The values of $R_{n-1}/R_n$ which will be required in Part II are given in Table I, as calculated from 7·12:

TABLE I.

| $\beta$ | $\frac{R_1}{R_2}$ | $\frac{R_2}{R_3}$ | $\beta$ | $\frac{R_1}{R_2}$ | $\frac{R_2}{R_3}$ |
|---|---|---|---|---|---|
| 7 | 1·42 + 1·17 $i$ | | 100 | 20·4 + 20 $i$ | 14·7 + 14·3 $i$ |
| 8 | 2·02 + 1·56 $i$ | | 150 | 30·4 + 30 $i$ | 21·9 + 21·5 $i$ |
| 9 | 2·23 + 1·76 $i$ | | 200 | 40·4 + 40 $i$ | 29·1 + 28·7 $i$ |
| 10 | 2·43 + 1·98 $i$ | 1·60 + 1·38 $i$ | 250 | 50·4 + 50 $i$ | 36·3 + 35·9 $i$ |
| 11 | 2·63 + 2·17 $i$ | 2·01 + 1·50 $i$ | 300 | 60·4 + 60 $i$ | 43·5 + 43·1 $i$ |
| 12 | 2·82 + 2·37 $i$ | 2·38 + 1·71 $i$ | 350 | 70·4 + 70 $i$ | 50·7 + 50·3 $i$ |
| 15 | 3·41 + 2·98 $i$ | 2·61 + 1·97 $i$ | 400 | 80·4 + 80 $i$ | 58·0 + 57·5 $i$ |
| 20 | 4·41 + 3·96 $i$ | 3·31 + 2·84 $i$ | 450 | 90·4 + 90 $i$ | 65·3 + 64·9 $i$ |
| 30 | 6·41 + 5·96 $i$ | 4·73 + 4·27 $i$ | 500 | 100 + 100 $i$ | 72·5 + 71·8 $i$ |
| 40 | 8·41 + 7·96 $i$ | 6·16 + 5·71 $i$ | 550 | 110 + 110 $i$ | 79·6 + 79·0 $i$ |
| 50 | 10·4 + 9·96 $i$ | 7·60 + 7·16 $i$ | 600 | 120 + 120 $i$ | 86·7 + 86·0 $i$ |
| 60 | 12·4 + 12·0 $i$ | 9·02 + 8·58 $i$ | 700 | 140 + 140 $i$ | 101 + 100 $i$ |
| 70 | 14·4 + 14 $i$ | 10·4 + 10·0 $i$ | 800 | 160 + 160 $i$ | 115 + 114 $i$ |
| 80 | 16·4 + 16 $i$ | 11·9 + 11·4 $i$ | 900 | 180 + 180 $i$ | 129 + 128·5 $i$ |
| 90 | 18·4 + 18 $i$ | 13·3 + 12·8 $i$ | 1000 | 200 + 200 $i$ | 143 + 143 $i$ |

§ 8. *Induction in a conducting shell enclosing a conducting core separated from it by a non-conducting shell.*

The shell now to be considered will be supposed to have a magnetic permeability $\mu_s$, conductivity $\kappa$, and external and internal radii $a$, $qa$ respectively. The non-conducting material on the outside and inside will be assumed to have magnetic permeabilities $\mu_o$, $\mu_i$ respectively. The vector potential inside the conducting shell will have terms of both external and internal origin, the former corresponding to the primary external field, the latter to the induced currents in the central core.

The forces in the non-conducting material inside and outside the shell can be deduced from the magnetic potentials, and may be written

$$B_r = -\mu\,\{nE\,(r/a)^{n-1} - (n+1)\,I\,(a/r)^{n+2}\}\,P_n^p e^{\iota p\phi} e^{at}, \quad\ldots\ldots 8{\cdot}1$$

$$H_\phi = -\iota p \operatorname{cosec}\theta\,\{E\,(r/a)^{n-1} + I\,(a/r)^{n+2}\}\,P_n^p e^{\iota p\phi} e^{at}, \quad\ldots\ldots 8{\cdot}2$$

$E$, $I$ and $\mu$ having different values, of course, on the two sides of the conducting shell.

In the substance of the shell **B** must be derived from the vector potential (cf. 5·8) and the $r$ and $\phi$ components are

$$B_r = -n(n+1)\,\{E_s\,(r/a)^{n-1} R_n + I_s\,(a/r)^{n+2} R_{-n-1}\}\,P_n^p e^{\iota p\phi} e^{at}, \quad\ldots\ldots 8{\cdot}3$$

$$B_\phi = -\frac{\iota p a^2}{r\sin\theta}\left[E_s\frac{\partial}{\partial r}\left\{\left(\frac{r}{a}\right)^{n+1} R_n\right\} + I_s\frac{\partial}{\partial r}\left\{\left(\frac{a}{r}\right)^n R_{-n-1}\right\}\right] P_n^p e^{\iota p\phi} e^{at}$$

$$= -\iota p \operatorname{cosec}\theta\left\{\left(\frac{r}{a}\right)^{n-1} E_s\rho_n(kr) - \left(\frac{a}{r}\right)^{n+2} I_s\rho_n'(kr)\right\} P_n^p e^{\iota p\phi} e^{at} \quad\ldots\ldots 8{\cdot}4$$

(by 6·4, 6·12), where

$$\rho_n(kr) = (n+1)\,R_n + (2n+3)^{-1} k^2 r^2 R_{n+1}, \quad\ldots\ldots 8{\cdot}5$$

$$\rho_n'(kr) = nR_{-n-1} + (2n-1)^{-1} k^2 r^2 R_{-n}. \quad\ldots\ldots 8{\cdot}6$$

Equating the normal induction and tangential force at the two sides of the inner and outer boundaries ($r = a$, $qa$) and eliminating $I_s$, $E_s$, the following expressions in terms of $I_o$ and $E_o$ are obtained for $I_i$ and $E_i$:

$$I_i = [\{q^{2n+1} F_o(ka)\,G_i(kqa) - G_o(ka)\,F_i(kqa)\}\,E_o + \{q^{2n+1} f_o(ka)\,G_i(kqa) + g_o(ka)\,F_i(kqa)\}\,I_o] \div H(ka), \quad\ldots 8{\cdot}7$$

$$E_i = [\{F_o(ka)\,g_i(kqa) + q^{-2n-1} G_o(ka)\,f_i(kqa)\}\,E_o + \{f_o(ka)\,g_i(kqa) - q^{-2n-1} g_o(ka)\,f_i(kqa)\}\,I_o] \div H(ka), \quad\ldots 8{\cdot}8$$

where

$$f(x) = n(n+1)\,\mu_s R_{-n-1} - (n+1)\,\mu\rho_n'(x), \quad\ldots\ldots 8{\cdot}9$$

$$g(x) = n(n+1)\,\mu_s R_n + (n+1)\,\mu\rho_n(x), \quad\ldots\ldots 8{\cdot}10$$

$$F(x) = n(n+1)\,\mu_s R_{-n-1} + n\mu\rho_n'(x), \quad\ldots\ldots 8{\cdot}11$$

$$G(x) = -n(n+1)\,\mu_s R_n + n\mu\rho_n(x), \quad\ldots\ldots 8{\cdot}12$$

the suffix $i$ or $o$ referring to the second $\mu$ in these expressions, while

$$H(x) = \mu_i\mu_s n(n+1)(2n+1)\,\{R_n\rho_n'(x) + R_{-n-1}\rho_n(x)\}. \quad\ldots\ldots 8{\cdot}13$$

If $\mu_o = \mu_i = 1$, and $q = 1 - \delta$ where $\delta$ is a small fraction, of which squares and higher powers are negligible, the above expressions for $I_i$ and $E_i$ lead to the results

$$(I_o - I_i)/\delta = [I_o\,\{(2n+1)\,fG + ka\,(fG' + gF')\} + E_o\,\{(2n+1)\,FG + ka\,(FG' - GF')\}] \div H, \ldots 8{\cdot}14$$

$$(E_o - E_i)/\delta = [I_o\,\{(2n+1)\,fg + ka\,(fg' - gf')\} + E_o\,\{-(2n+1)\,fG + ka\,(Fg' + Gf')\}] \div H, \ldots 8{\cdot}15$$

where the argument throughout on the left-hand side is $ka$, and the accent ($'$) denotes differentiation with respect to this argument.

If further $\mu_s = 1$, these equations, with the aid of 6·4, 6·12, 7·10 and some rather tedious algebra, reduce to

$$I_i = I_o + \frac{2\iota\beta^2\delta}{2n+1}\left(I_o - \frac{n}{n+1}E_o\right), \quad \text{8·16}$$

$$E_i = E_o + \frac{2\iota\beta^2\delta}{2n+1}\left(\frac{n+1}{n}I_o - E_o\right). \quad \text{8·17}$$

§ 9. *Earth currents, or earth potential gradients.*

The surface values of the potential gradients which impel the induced currents within the conducting earth or conducting shell can be determined as follows. The vector potential is continuous across the boundary of the conductor, so that just beneath the surface

$$A_r = 0, \quad \text{9·1}$$

$$A_\theta = -\frac{\iota p a}{\sin\theta}\left\{\frac{E_o}{n+1} - \frac{I_o}{n}\right\} P_n^p e^{\iota p\phi} e^{\alpha t}, \quad \text{9·2}$$

$$A_\phi = a\left\{\frac{E_o}{n+1} - \frac{I_o}{n}\right\}\left(\frac{\partial}{\partial\theta}P_n^p\right) e^{\iota p\phi} e^{\alpha t}. \quad \text{9·3}$$

The corresponding values of the components of the electric force just beneath the surface, in electromagnetic units, are (cf. § 3)

$$E_r = 0, \quad \text{9·4}$$

$$E_\theta = \frac{\iota p \alpha a}{\sin\theta}\left\{\frac{E_o}{n+1} - \frac{I_o}{n}\right\} P_n^p e^{\iota p\phi} e^{\alpha t}, \quad \text{9·5}$$

$$E_\phi = -\alpha a\left\{\frac{E_o}{n+1} - \frac{I_o}{n}\right\}\left(\frac{\partial}{\partial\theta}P_n^p\right) e^{\iota p\phi} e^{\alpha t}. \quad \text{9·6}$$

Further, since

$$H_r = -\{nE_o - (n+1)I_o\} P_n^p e^{\iota p\phi} e^{\alpha t} \quad \text{9·7}$$

and $a = 2.10^9/\pi$ at the earth's surface, it follows that

$$E_\theta = -2\iota p\alpha.10^9 H_r/\{n(n+1)\pi\sin\theta\}. \quad \text{9·8}$$

Expressed as a potential gradient in volts per kilometre per $1\gamma$ ($10^{-5}$ C.G.S.) of $H_r$ (the coefficient in the expression for the vertical magnetic force), this is

$$-\frac{20\iota p\alpha}{n(n+1)\pi\sin\theta} \quad \text{9·9}$$

The corresponding coefficient in the latitudinal component $E_\phi$ of the potential gradient is

$$\frac{20\alpha}{n(n+1)\pi}\cdot\left(\frac{\partial}{\partial\theta}P_n^p\right)\Big/P_n^p. \quad \text{9·10}$$

## PART II

### *Applications to problems of terrestrial magnetism.*

§ 10. *The influence of the permeability of the core.*

Though it is unlikely that the permeability $\mu_s$ differs much from unity in the inner part of the earth, it is not without interest to consider how the estimates of the conductivity $\kappa$ and the radius $qa$ of the supposed conducting core of the earth would be affected if a larger value of $\mu$ were assumed. The estimates will be supposed based on the amplitude-ratio and phase-difference between the surface values of the potentials of the external and internal portions of the field of the diurnal magnetic variations, and the values of these data determined by S. Chapman will be adopted, viz.*

$$\text{Amplitude-ratio (external : internal)} = 2{\cdot}55 : 1,$$
$$\text{Phase-difference (external} - \text{internal)} = -19^\circ.$$

In the notation of Part I this signifies that

$$I_o/E_o = 0{\cdot}371 + 0{\cdot}128i.$$

The value found in § 7 for this ratio $I_o/E_o$ is given by 7·8, where, for any given harmonic constituent in the potential, the unknown quantities on the right are $\mu_s$, $q$, and $\kappa$ (involved in $kqa$, the argument of the functions $R_n$, $R_{n+1}$). In the memoir cited it was assumed that $\mu_s = 1$; the results obtained by assuming other values instead of this are given in the following table, for the two principal harmonics $P_2^1$ and $P_3^2$ in the magnetic potential. The results are expressed in terms of $\kappa$ (in electromagnetic units) and $d$, the depth, in miles, of the supposed non-conducting outer shell of the earth ($d = (1-q)a$ expressed in miles). The different values of $d$ and $\kappa$ obtained

TABLE II.

| $\mu$ | $P_2^1$ | | $P_3^2$ | |
|---|---|---|---|---|
| | $d$ | $\kappa$ | $d$ | $\kappa$ |
| 1 | 168 | $3{\cdot}45 \,.\, 10^{-13}$ | 176 | $3{\cdot}31 \,.\, 10^{-13}$ |
| 10 | 180 | $3{\cdot}63 \,.\, 10^{-12}$ | 188 | $3{\cdot}39 \,.\, 10^{-12}$ |
| 100 | 180 | $3{\cdot}54 \,.\, 10^{-11}$ | 188 | $3{\cdot}33 \,.\, 10^{-11}$ |

from the two harmonics merely indicate that the slightly different observed values of $I_o/E_o$ for the two harmonics should be taken into account; this was not done because the difference between them is within the margin of error of the determinations.

It appears from the above table that it is mathematically possible for $\mu$ to have a wide range of values, consistently with the observed values of $I_o/E_o$, and without much affecting the estimate of $d$; the estimate of $\kappa$ varies almost proportionately with the value assumed for $\mu$.

§ 11. *The influence of the surface layers of ocean or moist earth.*

In the calculations of § 10 the existence of the conducting layers of ocean and moist earth near the earth's surface has been ignored. But the conductivity of sea-water (here taken as $4 \,.\, 10^{-11}$) much exceeds that of the core as hitherto estimated, so that it is desirable to examine how far the shallow oceanic layers can affect the ratio of the external and internal fields, and the estimates

* *Phil. Trans.* A 218, p. 1, 1919.

of the core. In so doing, the surface conducting layer will for simplicity be supposed uniform and to envelope the earth completely. The depths considered will not exceed 5 miles, so that the quantity $\delta$ of § 8 is less than 1/800, and the approximations made in obtaining 8·16, 8·17 are legitimate.

Table III shows the values of $I_i/E_i$ calculated for the undersurface of various depths of shallow ocean corresponding to the value $0{\cdot}371 + 0{\cdot}128i$ for $I_o/E_o$ as already used in § 10. The corresponding estimates of $\kappa$, the conductivity, and $d$, the depth of the surface of the core, are also given in most of the cases.

TABLE III.

| Depth of sea-water | $P_2^1$ | | | $P_3^2$ | | |
|---|---|---|---|---|---|---|
| | $I_i/E_i$ | $d$ | $\kappa$ | $I_i/E_i$ | $d$ | $\kappa$ |
| 125 feet | $0{\cdot}371 + 0{\cdot}126\,i$ | 168 | $3{\cdot}55\,.\,10^{-13}$ | $0{\cdot}367 + 0{\cdot}123\,i$ | 191 | $5{\cdot}7\,.\,10^{-13}$ |
| 250 feet | $0{\cdot}368 + 0{\cdot}123\,i$ | 178 | $3{\cdot}5\,.\,10^{-13}$ | $0{\cdot}363 + 0{\cdot}116\,i$ | 210 | $6{\cdot}9\,.\,10^{-13}$ |
| $\frac{1}{4}$ mile | $0{\cdot}349 + 0{\cdot}109\,i$ | 198 | $4{\cdot}7\,.\,10^{-13}$ | $0{\cdot}340 + 0{\cdot}058\,i$ | 336 | $13{\cdot}9\,.\,10^{-13}$ |
| $\frac{1}{2}$ mile | $0{\cdot}343 + 0{\cdot}071\,i$ | 340 | $10{\cdot}4\,.\,10^{-13}$ | $0{\cdot}328 + 0{\cdot}008\,i$ | — | — |

It thus appears that even a comparatively shallow sea ($\frac{1}{4}$ or $\frac{1}{2}$ mile in depth) affects the estimates of the core very considerably, and that the influence is much the greater on the component of higher degree ($P_3^2$). A shell of sea-water of depth much exceeding half a mile would cause the imaginary part of $I_i/E_i$ to become negative*, which is incompatible with any possible size and conductivity of the core. This suggests that the ratio adopted for $I_o/E_o$ cannot be valid for stations in or near deep seas. The question may be otherwise illustrated by calculating the ratio of $I_o/E_o$ corresponding to a core of size and conductivity as estimated in S. Chapman's memoir, together with an oceanic layer (assuming various alternative depths for this latter), the primary magnetic field being of purely external origin. The values of $E_o/I_o$ and of the phase-difference are given in Table IV, with similar comparative data also for the external parts of the field just inside and just above the oceanic layer.

TABLE IV.

*Sphere with core as estimated, surrounded by ocean.*

| Depth of ocean in miles | Component $P_2^1$ | | | | Component $P_3^2$ | | | |
|---|---|---|---|---|---|---|---|---|
| | $E_o/I_o$ | | $E_i/E_o$ | | $E_o/I_o$ | | $E_i/E_o$ | |
| | Amp. | Phase | Amp. | Phase | Amp. | Phase | Amp. | Phase |
| 0 | 2·49 | 19° | 1 | 0° | 2·41 | 19° | 1 | 0° |
| $\frac{1}{2}$ | 2·18 | 20 | ·917 | 9 | 2·00 | 21 | ·885 | 14 |
| 1 | 1·96 | 22 | ·835 | 15 | 1·72 | 20 | ·763 | 25 |
| 2 | 1·75 | 18 | ·678 | 27 | 1·51 | 16 | ·568 | 38 |
| 3 | 1·66 | 15 | ·556 | 35 | 1·43 | 13 | ·441 | 46 |
| 4 | 1·59 | 13 | ·472 | 40 | 1·40 | 10 | ·351 | 51 |

In columns 3 and 7 of the above table the phase of $I_o$ is in advance, and in columns 5 and 9 the leading phase is that of $E_o$.

* Implying that the phase of $E_i$ is in advance of the phase of $I_i$.

In the case of both components the amplitude-ratio $E_o/I_o$ diminishes steadily, though not proportionately, as the ocean depth increases, while the phase-difference first increases and then diminishes. Clearly the mean potential deduced from a number of land and sea stations is not likely to represent the potential in a model earth consisting of a uniform core surrounded by a uniform ocean. Some difference should be observable between the potential as derived separately from land and sea stations, though in the case of the actual earth the exact difference cannot easily be calculated on account of the irregular distribution of land and sea.

The data given by S. Chapman have been examined from this standpoint, and though inadequate for a proper discussion of so complicated a matter, seem to lend support to the above conclusions. Table V shows the means of the amplitude-ratios and phase-differences separately calculated (by the method described on p. 20 of his memoir) from the data for the six continental stations there dealt with: the components considered are $P_3^2$ and $P_4^3$, $P_2^1$ being left out of account because of an irregularity noted in the paper cited. Particulars of the six stations are given in the first part of Table V.

TABLE V.

A. Particulars of continental stations.

| | Observatory | Lat. | $\theta$ | Long. (East +$^{ve}$) |
|---|---|---|---|---|
| 1 | Ekaterinburg ... | 56° 50′ | 33° 10′ | 60° 38′ |
| 2 | Potsdam ......... | 52° 23′ | 37° 37′ | 13° 4′ |
| 3 | Irkutsk ......... | 52° 16′ | 37° 44′ | 104° 19′ |
| 4 | Tiflis ............ | 41° 43′ | 48° 17′ | 44° 48′ |
| 5 | Baldwin ......... | 38° 47′ | 51° 13′ | −95° 10′ |
| 6 | Pilar ............ | −31° 41′ | 121° 41′ | −63° 51′ |

B. Amplitude-ratio and phase-difference observed.

| | $P_3^2$ | $P_4^3$ |
|---|---|---|
| Mean amplitude-ratio ......................... | 2·31 | 2·2 |
| Mean phase-difference ......................... | 24°·2 | 30° |

C. Amplitude-ratio and phase-difference calculated for a sphere of radius 4000 miles, consisting of a non-conducting shell 50 miles thick, enclosing a core of conductivity $1{\cdot}8\,.\,10^{-13}$ C.G.S. units.

| | $P_3^2$ | $P_4^3$ |
|---|---|---|
| Amplitude-ratio .............................. | 2·26 | 2·24 |
| Phase-difference .............................. | 25°·9 | 26°·5 |

The values of the amplitude-ratios and phase-differences deduced from the observations are compared in Table V, B and C with those calculated on the hypothesis of a core of conductivity $1{\cdot}8\,.\,10^{-13}$ with its surface at a depth of 50 miles, these being the constants giving the best fit with the observed values in Table V B (when no surface conducting layer is taken into account). The value of the conductivity of the core is somewhat less than that obtained from land and sea stations taken together, when the effect of the ocean is ignored. With such a core the currents resulting from the 24-hour term in the potential would have one-sixth of their surface value at a depth of 1500 miles, and for quicker periods at smaller depths.

An attempt was made to discover whether the effect of the ocean resembled that calculated in Table IV, by considering the data from stations surrounded by large and deep expanses of sea, viz. Honolulu, Batavia, and Christchurch (N.Z.). The averages of the amplitude-ratios for $P_3^2$ and $P_4^3$ were 1·8 and 1·6 respectively, i.e., considerably less (as was to be anticipated) than those for land stations. The phase-differences, however, were very diverse (e.g. for $P_3^2$ Honolulu gave 27°, and Christchurch 2°, Batavia being very different from either: while the irregularity was even more marked for $P_4^3$); this is perhaps not to be surprised at when it is considered how complicated the currents must be in an ocean of highly irregular contour and variable depth.

As regards the surface layer of moist earth, the conductivity of which is of the order $10^{-13}$, its influence—assuming any reasonable depth of the layer—is negligible compared with that of even a shallow sea, on account of the far greater conductivity of the water. Such a layer, even of so great a depth as ten miles, would only possess the same (total) conductivity as a layer of sea-water 125 feet deep. It is therefore reasonable to ignore the surface layer when considering data from continental stations far removed from great sheets of water.

§ 12. *The conductivity of the earth as deduced from magnetic storm data.*

The analysis of Part I will next be applied to the non-periodic magnetic variations observed during magnetic storms: in particular, to the "storm-time" portion*, representing the deviation of the field from the normal, averaged round the parallels of latitude. This part of the field being independent of the co-ordinate $\phi$, the $p$ of § 4 is zero. In middle latitudes the appropriate potential function is, to a first approximation,

$$\{r f_1(t) + (a^3/r^2) f_2(t)\} \cos\theta.$$

The easterly component is zero, while the northerly and vertical components are $\{f_1(t) + f_2(t)\} \sin\theta$ and $\{f_1(t) - 2f_2(t)\} \cos\theta$ respectively.

Curves showing the mode of variation of the horizontal and vertical components of the field are given in the paper just cited, and the typical curves for equatorial regions are reproduced here. The force in each case varies from zero to a maximum on one side and then reverses and attains a maximum on the opposite side, afterwards slowly recovering its original value. The simplest and most convenient mathematical expression capable of representing a function which increases from zero to a maximum and then decreases to zero asymptotically is $(e^{-lt} - e^{-l't})$, where $l' > l > 0$. The more complex expression $\Sigma A e^{-lt'}$, where $\Sigma A = 0$, suffices to represent the case where there are maxima on both sides of the normal. Thus if $f_1(t) = \Sigma E e^{-lt}$ and $f_2(t) = \Sigma I e^{-lt}$, where $\Sigma E = \Sigma I = 0$, the observed curves for the horizontal and vertical force can be represented. The values of $E$ and $I$ are not independent, however, but are connected by the relation 7·9 if the core (only) of the earth is taken into account: thus ($n$ being now unity)

$$\frac{I}{E} = \tfrac{1}{2} q^3 \left(1 - \frac{R_1}{R_c}\right) \qquad (kr = kqa).$$

* S. Chapman, "Outline of a Theory of Magnetic Storms," *Proc. Roy. Soc.* A 95, p. 61, 1918.

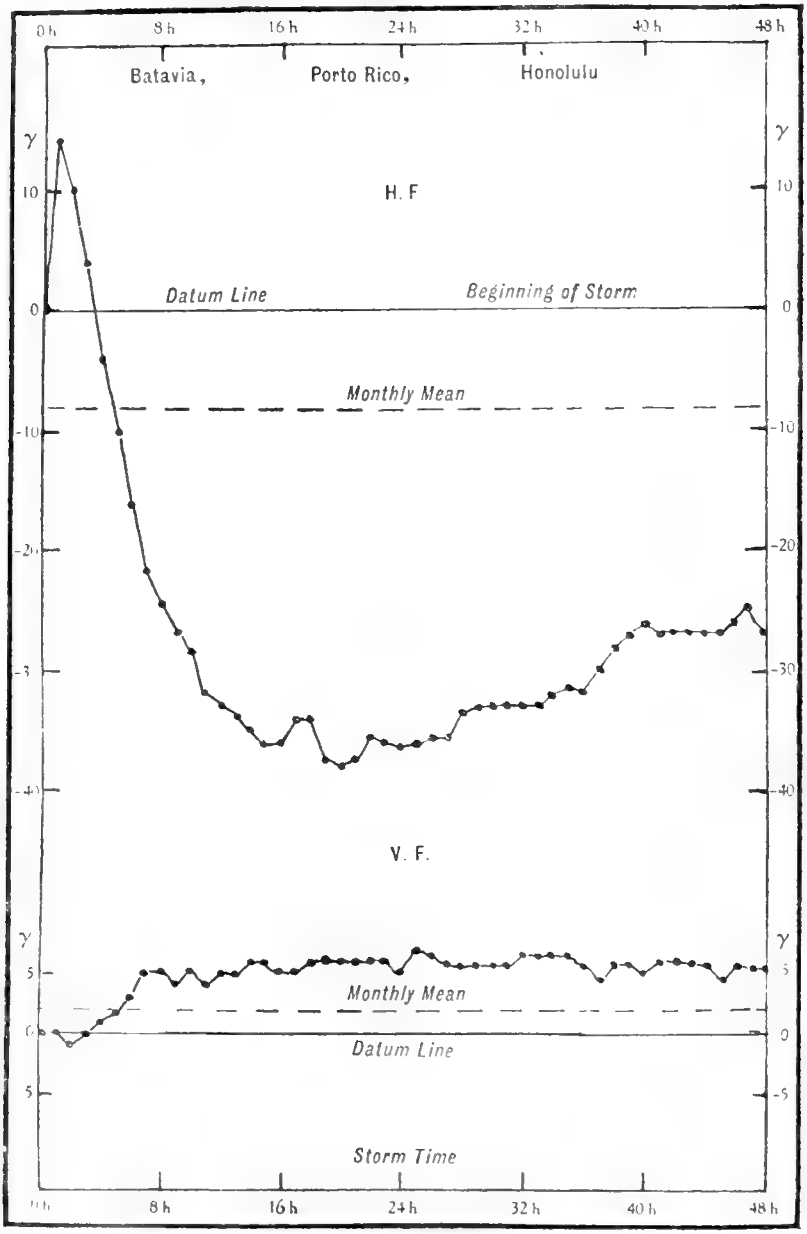



FIG. 1.

In the present case the $\alpha$ of Part I is negative, and equal to $-l$, so that (cf. 5·7) $k^2r^2$ is a real negative quantity, which will be written $-x^2$. The possible forms of $R_0$ are $\sin x/x$ and $\cos x/x$, but the latter is inappropriate here because it tends to infinity as $x$ (or $r$) tends to zero. The value of $R_1$ corresponding to the former value of $R_0$ is

$$-\frac{3}{x^2}\left(\cos x - \frac{\sin x}{x}\right),$$

so that

$$\frac{R_1}{R_0} = \frac{3}{x}\left(\frac{1}{x} - \cot x\right).$$

The values of $R_1/R_0$ corresponding to $x = m\pi + \theta$ are given in the following table.

TABLE VI.

| $\theta$ | $m=0$ | $m=1$ | $m=2$ | $m=3$ |
|---|---|---|---|---|
| 0 | + 1·00 | $-\ \infty$ | $-\ \infty$ | $-\ \infty$ |
| $\pi/10$ | 1·01 | − 2·43 | − 1·33 | − ·920 |
| $2\pi/10$ | 1·04 | − ·879 | − ·531 | − ·381 |
| $3\pi/10$ | 1·08 | − ·353 | − ·225 | − ·182 |
| $4\pi/10$ | 1·13 | − ·066 | − ·077 | − ·065 |
| $5\pi/10$ | 1·25 | + ·136 | + ·050 | + ·025 |
| $6\pi/10$ | 1·37 | ·315 | ·165 | ·110 |
| $7\pi/10$ | 1·62 | ·514 | ·299 | ·210 |
| $8\pi/10$ | 2·13 | ·826 | ·506 | ·368 |
| $9\pi/10$ | 3·66 | 1·64 | 1·05 | ·776 |
| $10\pi/10$ | $\infty$ | $\infty$ | $\infty$ | $\infty$ |

The ratio $R_1/R_0$ becomes infinite when $x = m\pi$, so that with a core of dimensions so related to the rate of decay as to make $x = m\pi$, the boundary conditions can be satisfied only if the field is entirely of internal origin. The external field could not itself decay at this rate in the presence of such a core.

Another interesting fact indicated by the above table is that $1 - \frac{R_1}{R_0}$ varies in sign as $x$ changes, so that the internal component of the field may reinforce either the horizontal or the vertical component of the external field: in the case of the diurnal magnetic variations it is the horizontal component which is reinforced, and the vertical component which is diminished.

In applying these considerations to the present case, the observed variations of the horizontal and vertical magnetic force were taken from the curves for the first (equatorial) group of stations, reproduced in Fig. 1 above from the paper cited. The simplest empirical formulae containing four negative exponentials—the same for the two curves—were then chosen to represent separately the initial and the main extremes on the two curves: the formulae were

$$\text{H.F.} = 40\,(e^{-6\cdot77t} - e^{-33\cdot2t}) - 125\,(e^{-0\cdot805t} - e^{-1\cdot831t}),$$

$$\text{V.F.} = -4\,(e^{-6\cdot77t} - e^{-33\cdot2t}) + 20\,(e^{-0\cdot805t} - e^{-1\cdot831t}),$$

$t$ being measured in days and the forces in $\gamma$. The first two terms in each case were chosen so as to fit the initial movement of the curves, and the second two terms to fit the main movement. The following table (VII) shows how nearly they represent the observed data: it may be added that the horizontal force deviation is reduced, according to this formula, to $1\gamma$ after 6 days.

While the agreement is good, the formulae are not theoretically correct, because the relations between the coefficients $E$ and $I$ in the expressions for $f_1(t)$ and $f_2(t)$ do not permit the coefficients to be equal and opposite in pairs (as above) *both* in H.F. and V.F. This unduly simple relationship must result in different values of the conductivity of the core being calculated from the comparison of corresponding coefficients in the formulae for the H.F. and V.F. It seems hardly worth while, however, to treat the present data more elaborately.

Owing to the rapidity of the first phase of the storm, the corresponding part of the variable field would probably not penetrate to the core of the earth for stations (like those here dealt with) near the ocean. Hence, and because of the comparatively small amplitude of the first phase

TABLE VII.

| Hour........... | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 9 | 12 | 18 | 24 | 36 | 48 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $40(e^{-6\cdot77t}-e^{-33\cdot27t})$ | 0 | 20 | 19 | 17 | 13 | 10 | 7 | 3 | 1 | 0 | 0 | 0 | 0 |
| $-125(e^{-0\cdot805t}-e^{-1\cdot831t})$ | 0 | −6 | −9 | −13 | −17 | −20 | −23 | −29 | −34 | −38 | −36 | −30 | −22 |
| H.F. (formula) ... | 0 | 14 | 10 | 4 | −4 | −10 | −16 | −26 | −33 | −38 | −36 | −30 | −22 |
| H.F. (observed) ... | 0 | 14 | 10 | 4 | −4 | −10 | −16 | −26 | −32 | −36 | −36 | −30 | −25 |
| $-4(e^{-6\cdot77t}-e^{-33\cdot27t})$ | 0 | −2 | −1·9 | −1·7 | −1·3 | −1 | −·7 | −·3 | −·1 | 0 | 0 | 0 | 0 |
| $20(e^{-0\cdot805t}-e^{-1\cdot831t})$ | 0 | 1·0 | 1·4 | 2·1 | 2·7 | 3·2 | 3·7 | 4·6 | 5·4 | 6·0 | 5·8 | 4·8 | 3·6 |
| V.F. (formula) ... | 0 | −1·0 | −·5 | ·4 | 1·4 | 2·2 | 3·0 | 4·3 | 5·3 | 6·0 | 5·8 | 4·8 | 3·6 |
| V.F. (observed) ... | 0 | 0 | −1 | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 6 | 5 | 5 |

of the storm, the calculation of the conductivity will refer only to the second phase represented by the second half of the above two formulae. The mean co-latitude $\theta$ of the group of stations is 68°, so that

$$E+I=-125 \operatorname{cosec} 68^\circ=-135,$$
$$E-2I=-20 \sec 68^\circ=-53^*,$$

so that

$$I/E=0\cdot26=\tfrac{1}{2}q^3\left(1-\frac{R_1}{R_0}\right).$$

If the whole earth is assumed uniformly conducting, $q=1$; if the first 160 miles of the earth's crust is non-conducting (cf. § 10), $q^3=0\cdot88$, so that $R_1/R_0$ is 0·48 or 0·41 in the two cases Now, the unit of time being one day,

$$x^2=-k^2r^2=\frac{4\pi\kappa l r^2}{86400c^2},$$

or

$$\frac{\kappa}{c^2}=\frac{1\cdot69\,x^2}{10^{14}}\cdot\frac{1}{l}.$$

Substituting the two values of $l$, 0·805 and 1·831, the following are the (necessarily different) values found for $\kappa\, c^2$, the value of the conductivity in electromagnetic units. The figures given are the minimum values, higher values being possible as indicated by Table VI.

TABLE VIII.

*Conductivity of the earth as deduced from storm data.*

| | $l=0\cdot805$ | $l=1\cdot831$ | Mean |
|---|---|---|---|
| No non-conducting crust ... | $5\cdot9\,.\,10^{-13}$ | $2\cdot6\,.\,10^{-13}$ | $4\cdot2\,.\,10^{-13}$ |
| Crust 160 miles deep ......... | $5\cdot6\,.\,10^{-13}$ | $2\cdot5\,.\,10^{-13}$ | $4\cdot1\,.\,10^{-13}$ |

The agreement of these estimates with that ($3\cdot65\,.\,10^{-13}$) obtained, using much more material, from the diurnal magnetic variations, would seem to be not unsatisfactory†. Doubtless with

* In the paper cited the V.F. is measured positive downwards, hence the sign in this equation.

† It may be seen from Table VI that any positive fractional value for $I/E$ leads to a minimum value for $\kappa$ of order $10^{-13}$.

some trouble four constants $A$ and four constants $B$ might be found, such that ($\Sigma A$ and $\Sigma B$ being zero) $\Sigma Ae^{-lt}$ and $\Sigma Be^{-lt}$ would represent the observed data of Table VII, and at the same time the corresponding $A$'s and $B$'s would be consistent with a single value of the conductivity of the core; this will not be attempted, however, since the diurnal variations are better adapted for giving an accurate and unambiguous estimate of the conductivity. The above discussion at least suggests that the estimate of $\kappa$ already obtained is not inconsistent with the magnetic storm variations.

§ 13. *Earth currents, or earth potential gradients.*

The earth currents to be considered here are of two kinds, viz., (i) those associated with the diurnal magnetic variations, small in magnitude but world-wide in distribution, and (ii) those associated with magnetic disturbance, particularly those which are local and of short duration.

TABLE IX.

*Observed and calculated values of southerly diurnal earth currents.*

| Hour | Actual E.M.F. | | | Reduced E.M.F. | |
|---|---|---|---|---|---|
| | Observed | | Calculated (1902 and 1905 mean) | Observed B | Calculated |
| | A | B | | | |
| Noon | 227 | 648 | 241 | 96 | 100 |
| 1 | 262 | 395 | 188 | 59 | 78 |
| 2 | 232 | 46 | 63 | 6 | 26 |
| 3 | 187 | −237 | −61 | −35 | −25 |
| 4 | 187 | −368 | −125 | −55 | −52 |
| 5 | 198 | −363 | −120 | −54 | −50 |
| 6 | 151 | −287 | −82 | −42 | −34 |
| 7 | 60 | −216 | −50 | −32 | −21 |
| 8 | −19 | −173 | −42 | −26 | −17 |
| 9 | −119 | −153 | −37 | −22 | −15 |
| 10 | −155 | −125 | −21 | −19 | −9 |
| 11 | −152 | −79 | 5 | −12 | 2 |
| 12 | −166 | −43 | 23 | −7 | 9 |
| 13 | −102 | −20 | 23 | −3 | 9 |
| 14 | −72 | −9 | 12 | −1 | 5 |
| 15 | −44 | −1 | 6 | 0 | 2 |
| 16 | −14 | −17 | 5 | −2 | 2 |
| 17 | −28 | −74 | −6 | −11 | −3 |
| 18 | −71 | −136 | −42 | −20 | −17 |
| 19 | −142 | −144 | −93 | −21 | −39 |
| 20 | −207 | −28 | −102 | −4 | −42 |
| 21 | −195 | 212 | −47 | 32 | −20 |
| 22 | −67 | 494 | 70 | 74 | 29 |
| 23 | 77 | 678 | 191 | 100 | 79 |

As regards (i), the formulae 9·9, 9·10 give the voltage impelling the currents, provided $\alpha$ is equated to $\iota p$, $p$ being the frequency of the component considered. If the value of the radial magnetic force $H_r$ is obtained for each harmonic constituent of the field, the diurnal variation in the voltage impelling the southerly earth currents can be at once determined.

Probably the most complete records of the diurnal variation of earth currents are those discussed by Weinstein, which depend on several years' records on lines from Berlin to Dresden and to Thorn*. Table IX, column B, gives the north-to-south component of the E.M.F. found from these earth current records, while column A contains the corresponding results derived from Airy's observations at Greenwich, made much earlier. The fourth column gives the diurnal variation of southerly E.M.F. calculated as above from the Fourier coefficients of the radial diurnal magnetic force variations for latitude 52 N. given at the end of S. Chapman's memoir on the diurnal magnetic variations, the mean value for 1902 and 1905 being adopted. The units in the first two columns of the table are arbitrary, and no relation between them is available. For the calculated value the unit is $10^{-5}$ volt per kilometre distance between the earth plates. The columns headed 'Reduced E.M.F.' have been added to show the agreement of type between column B and the calculated value, by reducing the maximum to 100 and keeping the same zero in each case.

These reduced values are also illustrated by the following graph:

DIURNAL EARTH CURRENTS.

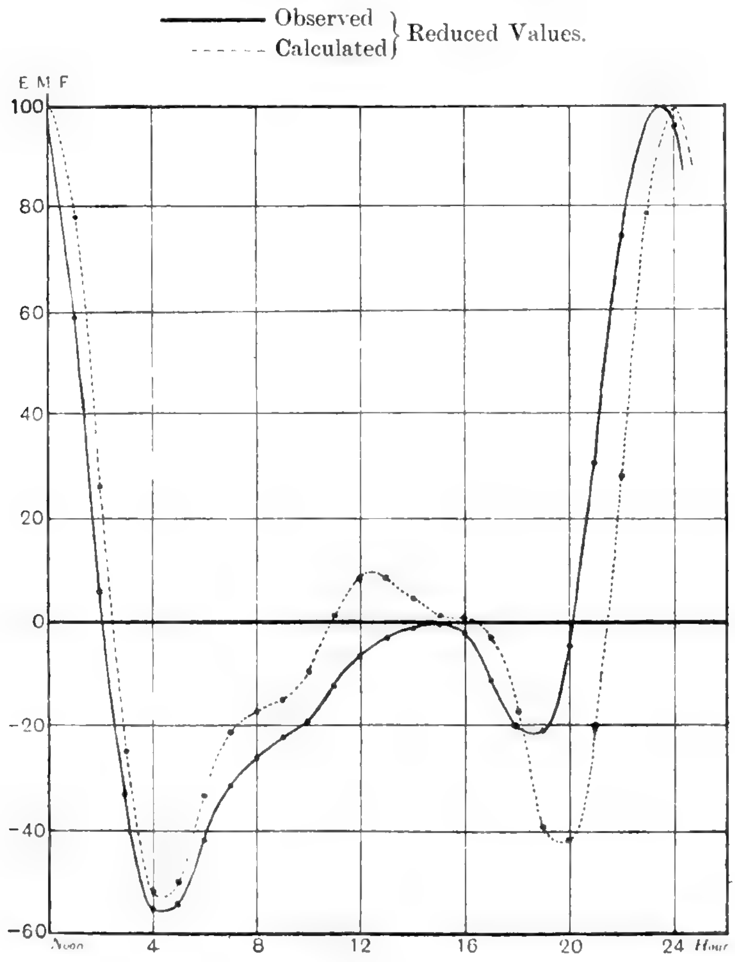



Fig. 2.

* Cf. C. Chree, article "Earth Currents" in the *Encyc. Brit.*

The calculated values agree better with Weinstein's values than with Airy's, which may be due either to better or more modern methods of observation or to the proximity of Greenwich to the sea. The agreement with Weinstein's data seems fairly satisfactory.

As regards the easterly E.M.F. of the diurnal earth currents, the agreement is less good. This is not unnatural, on account of the failure of the magnetic potential function to represent properly the northerly component of the magnetic force*, on which the easterly earth currents largely depend. The amplitude and character of the observed and calculated variation show very fair agreement, but there is a phase-difference of about $2\frac{1}{2}$ hours.

The local and irregular earth currents will next be considered. These are often large, e.g. a potential difference of between seven and eight hundred volts was found between earth plates 500 kilometres apart in 1859, and potential gradients of half a volt per mile have often been recorded since. The areas specially affected may vary within wide limits. The more local the area affected by a disturbance and its associated earth currents, the higher will be the degree and order of the leading terms when the field is analysed into its component spherical harmonics. The disturbance will be supposed oscillatory, so that in the time factor $e^{\alpha t}$ of Part I the appropriate value of $\alpha$ will be $\iota p$, where the frequency $p$ is supposed to be quite independent of the degree $n$ and order $m$ of the tesseral harmonics; $m$ and $n$ will be supposed fairly large, say from 5 to 10, in order to represent a disturbance completing its range over a relatively small fraction of the earth's surface. The period $2\pi/p$ may be anything from a few hours down to a few seconds.

The question will be illustrated chiefly in connection with the variations in the vertical magnetic force. Table X shows for latitude 60° N., and for a field varying in one harmonic component only, the amplitude of variation in the vertical force $H_r$ corresponding to an earth voltage of ·5 volt per mile; the variations are of course proportional to one another. The values of $n$ and $m$ refer to the degree and order of the few typical harmonic components considered, while the periods dealt with vary from 2 to 30 minutes.

TABLE X.

*Amplitude of $H_r$ when voltage reaches $\frac{1}{2}$ volt per mile.*

| $n$ | $m$ | Amplitude of $H_r$ in $\gamma$, with period in minutes | | | |
|---|---|---|---|---|---|
| | | 2 min. | 4 min. | 10 min. | 30 min. |
| 6 | 5 | 7·5 | 15·1 | 37·8 | 113 |
| 7 | 6 | 8·4 | 16·8 | 42 | 126 |
| 8 | 7 | 9·3 | 18·5 | 46·3 | 139 |
| 9 | 8 | 10·5 | 21·1 | 52·8 | 162 |
| 10 | 9 | 11 | 22 | 55 | 175 |

The corresponding variation of the horizontal force is less readily calculated, since it depends on the conductivity of the crust and core of the earth. With the data derived in § 11 from land

* Cf. *Phil. Trans.* A 218, p. 23, 1919.

stations only (immediately after Table V) the amplitudes of the resultant horizontal force variation corresponding to the first line of the last table above would be

| Period ....... | 2 min. | 4 min. | 10 min. | 30 min. |
|---|---|---|---|---|
| Amplitude in $\gamma$ | 203 | 290 | 446 | 892 |

Oscillations with so large an amplitude as $892\gamma$ in the horizontal force are rarely if ever observed, indicating that the surface potential gradient in the earth approaches half a volt per mile only when the period is shorter than 30 minutes.

# XXVI. *The Escape of Molecules from an Atmosphere, with special reference to the Boundary of a Gaseous Star.*

By E. A. Milne.

[*Communicated* by Prof. H. F. Newall, F.R.S., Director of the Solar Physics Observatory, Cambridge. *Received* 15 January, 1923.]

§ 1. *Scope of the paper.* It has recently been suggested* that since for a giant star of the observed size of α Orionis the value of gravity at the surface can be at most a fraction of that at the surface of the moon and since the moon is unable to retain an atmosphere, therefore the atmospheres of such stars would be rapidly dissipated, and consequently the stars themselves if assumed to be gaseous would be dissipated also. The fallacy in this argument arises from the fact that the rate of dissipation of an atmosphere depends not on the gravitational acceleration but on the gravitational potential, since it is the latter which determines the critical velocity of escape of the molecules. The gravitational potential falls off only as the inverse first power of the distance, and it is easily calculated that for α Orionis, in spite of its large radius and consequently low surface value of gravity, the potential at the surface is relatively large, large enough moreover for the loss by diffusion to be inappreciable. The point nevertheless suggests that it would be of interest to apply the detailed theory of the escape of molecules from an atmosphere to stars of various masses and temperatures.

In the case of the atmospheres of the earth, moon and planets the question has received considerable attention. Johnstone Stoney in 1868 pointed out that on the kinetic theory of gases a proportion of the molecules would from time to time attain speeds greater than the critical speed of escape from the gravitational pull of the planet, and that such molecules if moving outwards in the regions of low density where collisions are rare would be permanently lost to the atmosphere; and he afterwards elaborated this in a series of papers†. The subject has also been considered by Sir George Darwin‡, Cook§, Bryan‖ (chiefly from the point of view of the effects of rotation), Emden¶ and others, and a very clear summary of the method and the results of its application has been given by Jeans**. But none of these writers evaluates with any precision the height at which loss becomes appreciable, nor do they determine the critical density corresponding to this height, *i.e.* the density of the layer from which escape is mainly proceeding; again, one finds no definite estimate of the size of mean free path necessary in order that the chance of a collision may be sufficiently small. The reason appears to be that in certain cases precise knowledge of these quantities is immaterial. Jeans discusses in a general way the height of the critical level, but obtains his final result in a form independent of an evaluation of this height. For the rate of loss from an isothermal atmosphere he derives a formula expressed by the product of the critical density into a function of the critical height (quoted as formula (30)

* W. H. Pickering, *Pop. Astron.*, 28, no. 2, 1920; 29, no. 5, 1921.

† *Trans. Roy. Soc. Dublin*, 6, 305, 1898; *Astrophys. Journ.*, 1898—1904 (6 papers); *Proc. Roy. Soc.*, 67, 286, 1900.

‡ *Phil. Trans.*, 180, 1, 1889. ("On the mechanical condition of a swarm of meteorites.")

§ *Astrophys. Journ.*, 11, 36, 1900.

‖ *Brit. Assoc. Reports*, p. 682, 1893, p. 100, 1894, and p. 634, 1899; *Phil. Trans.*, 196 A, 1, 1900.

¶ *Gaskugeln*, p. 270 (1907).

** *Dynamical Theory of Gases*, 2nd edition, chap. 15, p. 357, 1916.

below), but when the critical density is eliminated by the introduction of the density* at a fixed reference point in the lower atmosphere (say at the base of the stratosphere in the case of the earth), the critical height is also almost eliminated. Thus the considerations employed by Jeans do not of themselves provide detailed information concerning the height and density of the escape layer. But it will be found that Jeans' formula applies only to the case of an isothermal atmosphere. If the atmosphere is not isothermal, the critical height does not eliminate itself, and Jeans' procedure alone does not lead to a determinate result.

Now it has been shown by the author† that (on certain assumptions) if the temperature distribution in the atmosphere is supposed described by a formula of the type

$$T = kr^{-n},$$

where $r$ is the distance from the centre of the star, $k$ is a constant and $n$ is a variable function of $r$, then $n$ must lie between 0 and 2; the limiting case $n=0$ corresponds to absorption only in the extreme ultra-violet, the limiting case $n=2$ to absorption only in the extreme infra-red, whilst $n=\frac{1}{2}$ gives the case when the gas absorbs uniformly throughout the spectrum. It appears desirable, therefore, to formulate the theory in a form applicable to such temperature distributions. This we shall attempt to do in Section II below. The kernel of the method is the use of the "available solid angle," which, at any level, is the empty solid angle into which molecules possessing sufficient velocities at that level are free to escape. It will appear that by this means a direct rough evaluation of the density and height of the escape layer is possible, and that the formula for the rate of escape becomes determinate. It will appear also that in some respects the isothermal case ($n=0$) is quite special; for example it is only in this case that the rate of escape is independent of the size of the molecules.

As a preliminary, the hydrostatics of an atmosphere in which the temperature distribution follows the law $T=kr^{-n}$, where $n$ is a constant, is investigated in Section I. It is found necessary to examine the behaviour of the pressure and density at great distances in some detail, in order to be able to overcome certain difficulties concerning the convergence of integrals which arise in Section II. The physical ideas underlying the main investigation in Section II are simple, and it is unfortunate that the mathematical analysis becomes somewhat complicated; I have been unable to see how to avoid these complications save at an undesirable sacrifice of rigour.

Section III, which deals with applications of the results, draws attention to a certain minimal property possessed by the values of the gravitational potential at the surfaces of the existing stars.

## I. *The hydrostatics of a gaseous gravitating atmosphere in which the temperature falls off as the inverse nth power of the distance from the centre of the attracting nucleus.*

§ 2. In this section it is proposed to consider the distribution of pressure and density in an atmosphere of the type stated, on the assumption that the gas is a continuous medium capable of indefinite tenuity, *i.e.* ignoring the molecular structure of the gas. Besides its application to the main problem of this paper the question possesses considerable mathematical interest of its own. The analogous problem for atmospheres in "polytropic" equilibrium has been investigated by Emden‡.

* The partial density of the constituent in question is implied, here and elsewhere.

† *Monthly Notices, R. A. S.*, 82, 368, 1922.

‡ Emden, *Gaskugeln*, chaps. 4, 5, 6, 9, 10 (Leipzig, 1907). The particular case of the isothermal gas-sphere ($n=0$) has been treated by Tait, Kelvin (*Phil. Mag.* v, 23, 287, 1887), Darwin (*loc. cit.*), G. W. Hill (*Annals of Mathematics*, 4, 19, 1888), Ritter (*Wild. Ann.* 16, 166, 1882) and very fully by Emden (*Gaskugeln*, chap. 9). See also Jeans, *Cosmogony and Stellar Dynamics* (1919), p. 146.

The temperature distribution is given by the formula

$$\frac{T}{T_0} = \left(\frac{r_0}{r}\right)^n \quad \text{.........................................(1)},$$

where $T$ is the temperature at distance $r$ and the suffix 0 refers to some convenient reference level near the limb. It will be found that in all cases the atmosphere extends to infinity, but the behaviour for large values of $r$ depends closely on the value of $n$. When $n > 1$, the total mass is finite; when $n < 1$, it is infinite, though the pressure and density still tend to zero at infinity; for the critical case, $n = 1$, the mass is found to be finite. This varying behaviour is reflected in the usual exponential formula used to represent the pressure and density throughout a limited range of $r$. The formula is obtained by neglecting the effect of the mass of the atmosphere itself on the value of gravity, and for $n < 1$ it gives a density which does not tend to zero as $r$ tends to infinity. But for $n < 1$ the effective increase of the mass of the attracting nucleus with increase of $r$, due to the included mass of the atmosphere, cannot be neglected if the pressure and density are required for large values of $r$. Within the range in which the density is appreciable, the mass is completely negligible in comparison with that of the nucleus; but the method we shall subsequently adopt to calculate the loss by diffusion requires the use of infinite integrals, and these integrals are not in a convergent form for $n < 1$ unless the mass of the atmosphere is taken into account.

§ 3. Let $p$ be the pressure, $\rho$ the density, $g$ the acceleration due to gravity, $V$ the gravitational potential (taken positively) at a distance $r$ from the centre; and let the suffix 0 denote the values of these quantities at a given distance $r_0$. Let $M(r)$ be the total mass inside the sphere of radius $r$ (including atmosphere and nucleus). Further let $G$ be the constant of gravitation, $R$ the gas-constant, $m$ the mass of a molecule of the gas. (Only one kind of molecule is assumed to be present.)

Put
$$q = \frac{mV}{RT} = \frac{GmM(r)}{RTr} = \frac{GmM(r)}{RT_0 r_0} \cdot \left(\frac{r}{r_0}\right)^{n-1} \quad \text{..........................(2)},$$

so that $q$ is the ratio of the gravitational potential energy of a molecule at any level to $\frac{2}{3}$ of its mean kinetic energy at that level. It is known that the value of $q$ controls the order of magnitude of the escape of molecules by diffusion, since the smaller $q$, the larger the mean velocity compared with that of escape. The behaviour of $q$ for large values of $r$ depends on the value of $n$.

The equation of hydrostatic equilibrium is

$$\frac{dp}{dr} = -g\rho \quad \text{.........................................(3)},$$

where
$$g = \frac{GM(r)}{r^2}, \quad \frac{p}{\rho} = \frac{RT}{m}.$$

Using (1) we find
$$\frac{1}{p}\frac{dp}{dr} = -\frac{GmM(r)}{RTr^2} = -\frac{GmM(r)}{RT_0 r_0^2}\left(\frac{r}{r_0}\right)^{n-2} = -\frac{q}{r} \quad \text{.......................(4)}.$$

If we assume that as $r$ tends to infinity the mass of the included atmosphere tends to a finite limit small compared with the mass of the nucleus, we may ignore the change in $g$ due to the mass of the atmosphere itself and take $M$ as being constant. We have then

$$\frac{1}{p}\frac{dp}{dr} = -q_0\left(\frac{r}{r_0}\right)^{n-2},$$

giving on integration
$$\frac{p}{p_0} = e^{-\frac{q_0}{n-1}\left(\frac{r^{n-1}}{r_0^{n-1}}-1\right)} \quad \text{.......................................(5)}.$$

If $n > 1$, this formula makes $p \to 0$ as $r \to \infty$, and moreover it makes the integral $\int^{\infty} \rho r^2 dr$ converge, so that $M(r)$ has a finite limit as $r \to \infty$, in accordance with the above assumption.

If $n = 1$, (5) must clearly be replaced by its limiting form

$$p/p_0 = (r_0/r)^{q_0} \quad \text{.......(6).}$$

This makes $p \to 0$ as $r \to \infty$, but it makes $\int^{\infty} \rho r^2 dr$ converge only if $q_0 > 4$. If $q_0 \leqslant 4$, (6) makes $M(r) \to \infty$, which contradicts the assumption, and hence if $n = 1$, and $q_0 \leqslant 4$, (6) is not valid.

If $n < 1$, (5) when put in the form

$$p/p_0 = e^{-\frac{q_0}{1-n}\left(1 - \frac{r_0^{1-n}}{r^{1-n}}\right)} \quad \text{.......(7)}$$

is seen to make $p$ tend to a finite non-zero limit as $r \to \infty$, and thus makes $M(r) \to \infty$. Hence (7) is not valid for large values of $r$ when $n < 1$.

If in (5) we put $n = 0$, we obtain

$$p/p_0 = \rho/\rho_0 = e^{-q_0(1 - r_0/r)} \quad \text{.......(8)}$$

the usual formula* for an isothermal atmosphere. As we have seen, this is invalid for large values of $r$, making $p/p_0$ and $\rho/\rho_0$ tend to $e^{-q_0}$. The lack of approximation is merely formal, but it makes (7) and (8) very inconvenient to work with in certain circumstances.

§ 4. We now investigate equation (4) more fully. On substituting for $M(r)$ from the relation

$$M(r) - M(r_0) = 4\pi \int_{r_0}^{r} \rho r^2 dr$$

and differentiating, we find

$$r^2 \frac{d^2(\log p)}{dr^2} + (2 - n) r \frac{d(\log p)}{dr} = -\lambda p r^{2n+2} \quad \text{.......(9),}$$

where

$$\lambda = \frac{4\pi G}{r_0^{2n}} \left(\frac{m}{RT_0}\right)^2.$$

Put

$$\lambda^{\frac{1}{2n+2}} r = s.$$

Then

$$s^2 \frac{d^2(\log p)}{ds^2} + (2 - n) s \frac{d(\log p)}{ds} = -p s^{2n+2} \quad \text{.......(10).}$$

This equation admits of the singular solution

$$p = \frac{(1-n)(2n+2)}{s^{2n+2}} = \frac{(1-n)(2n+2)}{\lambda r^{2n+2}} \quad \text{.......(11),}$$

which gives $p$ positive only when $n < 1$. We shall see that when $n < 1$, (11) is the asymptotic solution of (9) for large values of $r$ whatever the initial conditions $p_0$ and $M(r_0)$. It depends only on the scale of the distribution of temperature, namely on the value of the quantity

$$r^n T = r_0^n T_0,$$

which is constant throughout the atmosphere.

It is convenient to adopt procedure similar to that used by Emden for polytropic equilibrium. Perform the series of substitutions

$$p = e^{-v}, \quad s = e^{\theta}, \quad z = v - (2n + 2)\theta,$$

* *E.g.* Jeans, *Gases*, p. 354.

so that
$$ps^{2n+2} = e^{-z},$$
and then put
$$dz/d\theta = y.$$
We find
$$y\left[\frac{dy}{dz} - (n-1)\right] = e^{-z} + (n-1)(2n+2) \quad\text{.........................(12)},$$
an equation of the first order. Real solutions of this equation correspond to real, positive values of $p$. The general form of the integrals of (12) may be most easily seen by first drawing the curves
$$dy/dz = \text{const.},$$

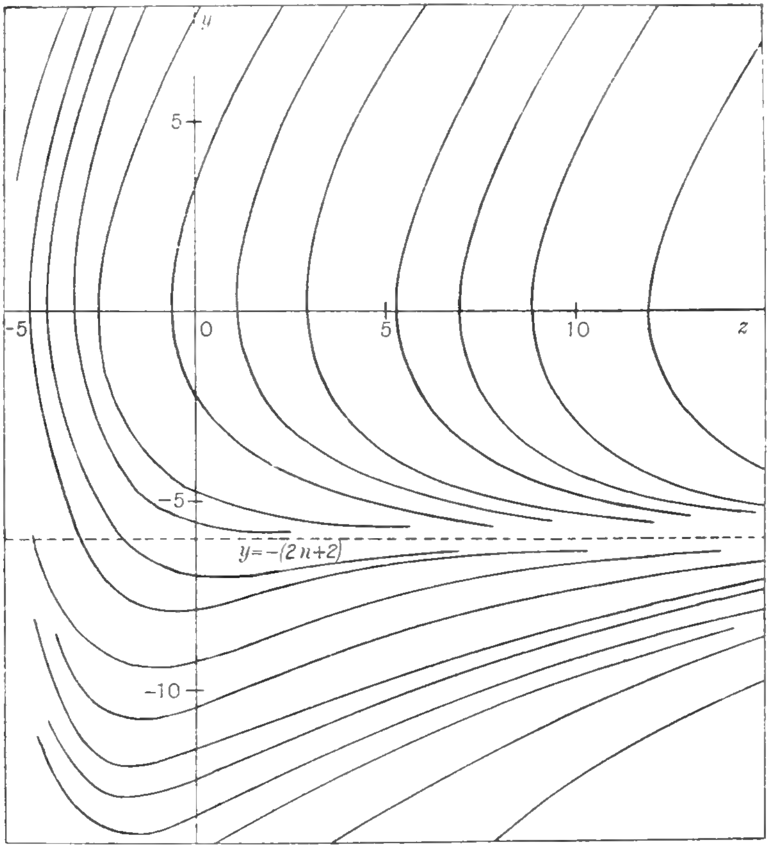



FIG. 1. Integrals of equation (12) for $n > 1$. (The curves are drawn for $n=2$.)

and marking each curve with a series of short transverse lines drawn in the direction given by the constant value of $dy/dz$. The actual integrals may then be rapidly sketched in by joining up the transversals belonging to neighbouring curves. All the curves pass through the point
$$y = 0, \quad z = -\log(1-n)(2n+2),$$
which corresponds to the singular solution.

When $n = 1$, (12) is integrable, and hence (10) must be integrable. We will consider this case first.

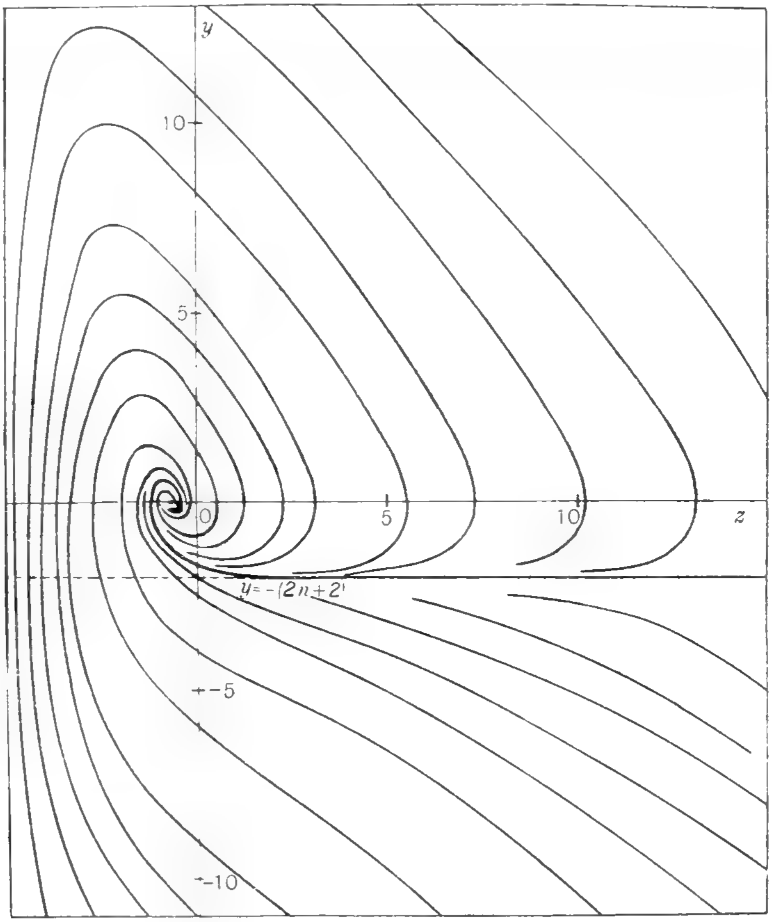



FIG. 2. Integrals of equation (12) for $n < 1$. (The curves are drawn for $n=0$.)

§ 5. *Case I.* $n=1$. It is found without trouble that for $n=1$ the solution of (10) is

$$p = \frac{1}{s^4} \cdot \frac{2A^2 B' s^A}{(B' s^A + 1)^2},$$

where $A$, $B'$ are arbitrary constants. In terms of $r$ this may be written

$$p = \frac{1}{\lambda r^4} \cdot \frac{2A^2 B (r/r_0)^A}{[1 + B (r/r_0)^A]^2} \quad \text{.......(13)},$$

where $B$ is a different arbitrary constant. The values of $A$ and $B$ are found to be, in terms of the initial conditions,

$$A^2 = (q_0 - 4)^2 + 2p_0 \lambda r_0^4 \quad \text{.......(14)},$$

$$B = \frac{A + (q_0 - 4)}{A - (q_0 - 4)} \quad \text{.......(15)}.$$

It can be verified that formula (13) reduces to the same value whether the positive or the negative value of $A$ be taken from (14). Accordingly we shall take the positive value. Equation (13) now shows that for large values of $r$,

$$p = O(r^{-(A+4)}),$$

$$\rho \propto p/T = O(r^{-(A+3)}),$$

and hence $\int^{\infty} \rho r^2 dr$ converges, since $A > 0$. We have thus proved that for $n = 1$ the mass is finite in all cases.

Equation (14) may be written

$$A^2 = (q_0 - 4)^2 + 2q_0^2 (4\pi p_0 r_0^2 / g_0 M_0) \qquad \text{(16).}$$

Now $4\pi p_0 r_0^2$ is the total pressure of outlying material outside the sphere of radius $r_0$, and $4\pi p_0 r_0^2/g_0$ is its mass, approximately. Thus $4\pi p_0 r_0^2/g_0 M_0$ is the ratio of the mass outside $r_0$ to the mass inside $r_0$. This is in general a very small fraction. Hence provided $q_0$ is not nearly equal to 4, we have approximately

$$A = |q_0 - 4| \qquad \text{(17).}$$

In practice (see Section III) $q_0$ is always large compared with 4, and so $(q_0 - 4)$ is positive; we find then approximately

$$B = 2(q_0 - 4)^2 / p_0 \lambda r_0^4,$$

which is large compared with unity. Substituting this value of $B$ in (13) and putting $A = q_0 - 4$, we find approximately

$$p/p_0 = (r_0/r)^{q_0},$$

in confirmation of (6). If, however, $q_0$ were less than 4, we should find for large values of $r$

$$p \propto r^{-4-(4-q_0)},$$

and (6) is no longer valid. In this case $4\pi p_0 r_0^2/g_0$ no longer measures the mass of the outlying atmosphere, the mass being in fact comparable with that of the nucleus. It is easy to show by direct integration that in all cases the mass of the atmosphere is given by

$$M(\infty) - M(r_0) = \frac{2AM(r_0)}{q_0(1+B)} \qquad \text{(18),}$$

and when $q_0$ is less than 4 the value of $B$ is given approximately by

$$B = \tfrac{1}{2} p_0 \lambda r_0^4 / (4 - q_0)^2,$$

which is *small* compared with unity; hence from (18) $M(\infty) - M(r_0)$ is in this case comparable with $M(r_0)$, as just stated.

Neglecting the variation of $M$ (*i.e.* when $q_0$ is large compared with 4) $q$ remains approximately constant, since

$$\frac{q}{q_0} = \frac{M(r)\, r_0 T_0}{M(r_0)\, r T} = 1.$$

*Case II.* $n > 1$. The case $n = 1$ appears to be the only one which is integrable in finite terms. The family of integrals of (12) is quite different according as $n \gtrless 1$. (See Figs. 1, 2.) When $n > 1$ the integrals have a common asymptote $y = -(2n+2)$, which they approach as $z \to +\infty$. In the neighbourhood of the asymptote they correspond to the region near the centre of the gas-sphere, which we are not here considering. Each integral also extends to infinity in the manner indicated by

$$y \sim (n-1)\, z, \quad (z \to +\infty),$$

and in this direction the curves correspond to the outside of the gas-sphere. When $z$ is large an approximate form of equation (12) is

$$y\left(\frac{dy}{dz} - (n-1)\right) = (n-1)(2n+2),$$

and by integrating this it can be shown that the corresponding asymptotic solution of the original equation (9) is

$$p \sim Ce^{-Dr^{n-1}} \qquad \text{.........(19)},$$

where $C$, $D$ are arbitrary constants. By comparing this with (9) it is found that $D = q_0/(n-1)$, provided the reference height $r_0$ is sufficiently large for (19) to be a valid approximation there. Thus (19) reduces to (5). It can also be shown that $p$ has no zero, so that there is no boundary for finite $r$. Neglecting the variation of $M$, the variation of $q$ is given by

$$\frac{q}{q_0} = \frac{M(r)\, r_0 T_0}{M(r_0)\, rT} = \left(\frac{r}{r_0}\right)^{n-1}.$$

*Case III.* $n < 1$. In this case it can be shown that the integrals of (12) are spiral curves whose convolutions approach the winding-point $z = -\log(1-n)(2n+2)$, $y = 0$, in a clockwise direction. In general each integral extends to infinity in the direction given by

$$y \sim -(1-n)\, z, \qquad (z \to +\infty)$$

which corresponds to the region near the centre of the gas-sphere. One particular integral, however, tends to infinity along the line $y = -(2n+2)$, and the convolutions of this integral about the winding-point separate the first convolutions of the other integrals from their second convolutions, their second convolutions from their third, and so on. The convolutions correspond to the outer parts of the gas-sphere. It can be shown, in fact, that as $z \to -\log(1-n)(2n+2)$ and $y \to 0$, $s \to \infty$. Further since $ps^{2n+2} = e^{-z}$, it follows that as $s \to \infty$,

$$ps^{2n+2} \to (1-n)(2n+2),$$

and hence the asymptotic solution of the original equation is

$$p \sim \frac{(1-n)(2n+2)}{\lambda r^{2n+2}} \qquad \text{.........(20)},$$

identical with the singular solution (11). Moreover as each spiral cuts the line $z = -\log(1-n)(2n+2)$ an infinite number of times, therefore there is an infinite number of places where $p$ is exactly equal to the expression on the right-hand side of (20). It can be shown that the true value of $p$ oscillates about that given by (20), the amplitude of the oscillations tending to 0 as $r$ tends to infinity. Further approximations can be obtained without trouble. It is worthy of note that the value of $p$ is ultimately independent of $p_0$ and $M(r_0)$.

From (20) it is easily deduced that for $r$ large,

$$\rho \sim \frac{(1-n)(2n+2)}{r^{n+2}} \frac{RT_0 r_0^{\,n}}{4\pi m G} \qquad \text{.........(21)},$$

$$M(r) \sim \int^r 4\pi\rho r^2 dr \sim \frac{(2n+2)\, RT_0 r_0}{mG} \left(\frac{r}{r_0}\right)^{1-n} \qquad \text{.........(22)},$$

whence

$$q = \frac{mGM(r)}{RTr} = \frac{mG}{RT_0} \frac{M(r)}{r} \left(\frac{r}{r_0}\right)^{n} \to 2n+2 \qquad \text{.........(23)}.$$

The differing behaviour of $q$ for $n \lessgtr 1$ is especially noteworthy.

In spite of the fundamental differences between the cases $n > 1$ and $n < 1$ *when $r$ is large*, it is clear that the exponential formula (5) which is always a good approximation to $p$ when $n > 1$ is also a good approximation to $p$ when $n < 1$ *within any limited range of $r$*. Can we obtain for $n < 1$ a single approximate formula for $p$ which will agree with (5) when $r/r_0$ does not depart far from unity and which will at the same time give the correct behaviour for large values of $r$?

Obviously to obtain such an approximation we must take account to some extent of the increase of mass. To do this let us solve equation (4) assuming that as far as the increase in mass is concerned $p$ may be represented by its asymptotic formula (20). In that case we have to insert in (4)

$$M(r) = M(r_0) + \frac{(2n+2)\,RT_0 r_0}{mG}\left(\frac{r}{r_0}\right)^{1-n} \qquad \text{.........................(24)},$$

and on integrating we obtain

$$\frac{p}{p_0} = \left(\frac{r_0}{r}\right)^{2n+2} e^{-\frac{q_0'}{1-n}\left(1-\frac{r_0^{1-n}}{r^{1-n}}\right)} \qquad \text{...........(25)},$$

where

$$q_0' = q_0 - (2n+2).$$

This combines the features of (5) and (20) in the required way. Moreover it agrees very closely with (5) for values of $r/r_0$ which depart little from unity. In fact the ratio of the value given by (25) to the more correct one given by (5) is

$$\left(\frac{r_0}{r}\right)^{2n+2} e^{\frac{2n+2}{1-n}\left(1-\frac{r_0^{1-n}}{r^{1-n}}\right)},$$

and if we put $r_0/r = 1-\delta$, to the second order this is equal to

$$1-(1-n^2)\,\delta^2.$$

Naturally (25) gives too small a value of $p$, since $M(r)$ does not increase so rapidly as is given by (24). But for our purpose (25) will prove most useful.

If in (25) we want to change the reference height from $r_0$ to $r_1$, it will be found that $q_0'$ must be replaced by a number $q_1'$, given by

$$q_1'/q_0' = (r_0/r_1)^{1-n}.$$

Hence we have the approximate formula for $q$,

$$\frac{q-(2n+2)}{q_0-(2n+2)} = \left(\frac{r_0}{r}\right)^{1-n} \qquad \text{.................................(26)};$$

this makes $q \rightarrow (2n+2)$ as $r \rightarrow \infty$, as we have already seen to be the case by a more rigorous method.

It may be asked which convolution of a spiral corresponds to the part of the atmosphere where the density is appreciable, near the limb. Now

$$y = -\frac{d\,(\log pr^{2n+2})}{d\,(\log r)} = -\frac{r\,dp}{p\,dr} - (2n+2) = q-(2n+2) \qquad \text{................(27)},$$

from (4), so that $y$ is initially positive and fairly large. Further

$$z = -\log \lambda p r^{2n+2} = -\log 4\pi G p \left(\frac{r}{r_0}\right)^{2n}\left(\frac{mr}{RT_0}\right)^2.$$

For a typical giant star for which $T_0 = 5100 \times 2^{-\frac{1}{4}} = 4300^\circ$, $r_0 = 8 \times 10^{11}$, taking $p_0 = 10^{-3}$ atmos., we find the initial value of $z$ for atomic hydrogen is about $+6$. Hence the point representing the pressure at the limb of a star is near the highest point of the first (outer) convolution of the corresponding spiral (Fig. 2). The decrease in pressure within a distance equal to the star's own radius corresponds to the upper part of the second half-sweep of the spiral about the winding-point.

## II. *The rate of escape of molecules from an atmosphere in which $T$ varies as $r^{-n}$.*

§ 6. *Introductory formulae.* The escape of molecules from an atmosphere will only begin to be appreciable (if at all) at a height such that above this height the density is so small that an outward-moving molecule has a reasonable chance of suffering no collision. Outward-moving molecules crossing this level with a velocity greater than the critical velocity $C$ given by

$$\tfrac{1}{2}mC^2 = mV$$

will possess enough energy for them to escape completely from the gravitational field of the nucleus, and will in fact so escape, describing hyperbolic orbits, provided they encounter no other molecules.

Let $r$ be a height so large that the chance of subsequent collision for a molecule crossing this level in an outward direction is negligible. Following JEANS we will calculate the total number of molecules crossing unit area per second with velocities exceeding $C$.

Let $dS$ be a small horizontal element of area surrounding the point $P$, which is situated at the height $r$. We will first calculate the number of molecules crossing $dS$ in time $dt$ with velocities lying between $c$ and $c+dc$. All such molecules must at the beginning of the interval $dt$ be lying inside the hemisphere below $dS$ of radius $cdt$ and centre $P$. Take a cone of small solid angle $d\omega$, with vertex at $P$ and axis making an angle $\theta$ with the vertical, and consider the element of volume $Q$ intercepted by it between hemispheres of radius $a$ and $a+da$, where $a < cdt$.

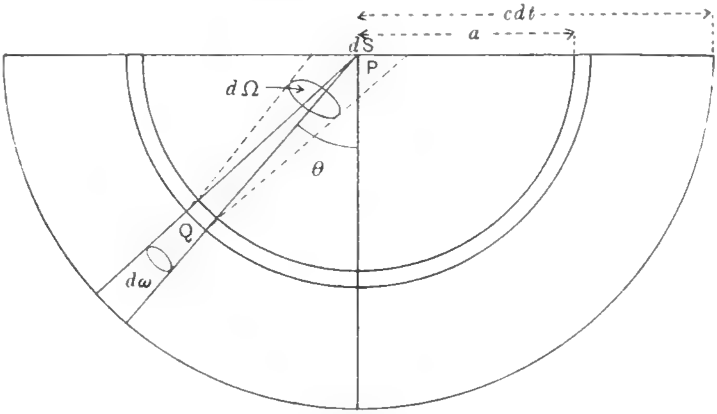



FIG. 3.

Provided $dt$ is small compared with the duration of a mean free path, all the molecules inside $Q$ whose velocities lie between $c$ and $c+dc$ and the directions of whose velocities lie inside the solid angle subtended by $dS$ at $Q$ will cross $dS$ during the interval $dt$.

If $\nu$ is the number of molecules per unit volume in the neighbourhood of $P$, the number per unit volume with velocities between $c$ and $c+dc$ and directions of velocities inside any given solid angle $d\Omega$ is, from the theory of gases,

$$\nu\left(\frac{m}{2\pi RT}\right)^{\frac{3}{2}} e^{-\frac{1}{2}mc^2/RT} c^2 dc\, d\Omega.$$

Hence the number of molecules inside $Q$ of the kind specified is, since here $d\Omega = dS\cos\theta/a^2$,

$$\nu\left(\frac{m}{2\pi RT}\right)^{\frac{3}{2}} e^{-\frac{1}{2}mc^2/RT} c^2 dc \frac{dS\cos\theta}{a^2} a^2 da\, d\omega.$$

Integrating with respect to $a$ from 0 to $cdt$, and integrating over the hemisphere, the total number crossing $dS$ with velocities between $c$ and $c+dc$ is

$$\nu\left(\frac{m}{2\pi RT}\right)^{\frac{3}{2}} dSdt \left(\int \cos\theta\, d\omega\right) e^{-\frac{1}{2}mc^2/RT} c^3 dc \quad\text{......................(28).}$$

Hence the total number crossing $dS$ with velocities exceeding $C$ is

$$\pi\nu\left(\frac{m}{2\pi RT}\right)^{\frac{3}{2}} dSdt \int_C^\infty e^{-\frac{1}{2}mc^2/RT} c^3 dc$$

$$= \nu\left(\frac{2RT}{\pi m}\right)^{\frac{1}{2}} dSdt\, \tfrac{1}{2} e^{-q}(1+q).$$

where, as in Section I, we have put

$$q = \frac{mV}{RT} = \frac{\frac{1}{2}mC^2}{RT} \quad\text{......................................(29).}$$

Hence the total *mass* of molecules crossing the sphere of radius $r$ per second with velocities exceeding the critical velocity is*

$$2\rho r^2\left(\frac{2\pi RT}{m}\right)^{\frac{1}{2}} e^{-q}(1+q) \quad\text{..................................(30).}$$

Having taken $r$ sufficiently large for outside collisions to be negligible, Jeans adopts this as giving the actual rate of loss.

Within a limited range of $r$, whatever the value of $n$, we have approximately

$$\frac{\rho}{\rho_0} = \frac{T_0}{T} e^{-\frac{q_0}{n-1}\left(\frac{r^{n-1}}{r_0^{n-1}}-1\right)} = \left(\frac{r}{r_0}\right)^n e^{\frac{q_0-q}{n-1}} \quad\text{..........................(31),}$$

since (within such a limited range)

$$q/q_0 = (r/r_0)^{n-1} \quad\text{........................................(32),}$$

the suffix 0 denoting as usual some convenient reference height in the lower atmosphere. Hence according to (30) the loss due to molecules crossing the sphere of radius $r$ is

$$2\rho_0 r_0^2\left(\frac{2\pi RT_0}{m}\right)^{\frac{1}{2}}\left(\frac{r}{r_0}\right)^{\frac{1}{2}n+2}(1+q)\, e^{\frac{q_0-nq}{1-n}} \quad\text{..........................(33).}$$

This is a function of $r$. But, on general grounds, although $r$ refers to a high level in the atmosphere the ratio $r/r_0$ will not be far from unity. The order of magnitude of (33) accordingly depends almost entirely upon the exponential factor.

Now when the atmosphere is isothermal, $n=0$ and the order of magnitude of (33) is independent of $r$. The rate of loss can therefore be evaluated without investigating the value of $r$, the height of the level from which escape may be said to be occurring, *i.e.* without investigating how rare collisions must be in order to be negligible. Further, as Jeans shows, the total mass of atmosphere lying above the level being approximately $4\pi r_0^2 p_0/g_0$, *i.e.* $4\pi r_0^2\rho_0 RT_0/mg_0$, the time to lose this amount of gas is, from (33) with $n=0$, equal to

$$\frac{1}{g_0}\left(\frac{2\pi RT_0}{m}\right)^{\frac{1}{2}}\left(\frac{r_0}{r}\right)^2 \frac{e^{-q_0}}{1+q} \quad\text{..................................(34),}$$

which, since in this we may put $r/r_0=1$, is independent of $\rho_0$. Thus the time needed for the streaming away of an amount of gas equal to the whole atmosphere above $r_0$ is independent of the total amount of gas above $r_0$.

* Jeans, *Gases*, chap. 15, p. 358. The notation has been changed slightly.

When $n \neq 0$, neither of these results is true. The rate of loss depends closely on $q$, and therefore on $r$. Hence it is necessary to investigate the level at which escape occurs, and further, since this height will depend on $\rho_0$, the time required for the escape of the outer atmosphere will depend on $\rho_0$.

We shall therefore analyse the mechanism of escape in more detail. Our analysis will furnish the so-far unknown orders of magnitude of the height $r$ and the density and mean free path at this height.

§ 7. *Outline of the method.* We shall begin by calculating the solid angle subtended at any point in the outer layers of the atmosphere by the molecules lying above that point. If we imagine an observer to ascend through the atmosphere from a level where the density is appreciable, and if we suppose further that all the molecules are visible to him as opaque bodies, then they will at first appear to fill his whole "sky." As he ascends he will in time reach a level such that the molecules in the zenith are just filling his sky without overlapping. As he ascends further, the sky in the region of the zenith will become partially clear; outside a certain zenith distance his sky will still be completely blocked, but within this zenith distance only a portion of his sky will be so covered at any one moment—there will be spaces visible between the molecules. The small circle separating the two regions may be called his molecular horizon. If he ascends further still, his molecular horizon will sink towards the horizontal, eventually sinking beneath it, and there will then be some clear sky in every outward direction.

Let $R$ be the height at which the zenith is just partially clearing. At this height some of the molecules moving *radially* outwards will be able to escape altogether, but molecules moving outwards in all other directions will suffer collisions. Hence the critical sphere must be taken outside $R$. Take now for $r$ the height at which the zenith distance of the molecular horizon is $\theta_1$. Take an elementary solid angle $d\omega$, the direction of whose axis has a zenith distance $\theta$, and let $f(r, \cos\theta)$ be the fraction of this solid angle which is actually occupied by molecules. The function $f(r, \cos\theta)$ is equal to unity for all zenith distances $\theta$ for which $\theta \geqslant \theta_1$. The solid angle available for molecules escaping through $d\omega$ is

$$[1 - f(r, \cos\theta)]\, d\omega.$$

Now formulae (28) and (30) were found on the assumption that the whole sky of solid angle $2\pi$ was free. It follows that, assuming approximately rectilinear* paths for the escaping molecules, the true rate of escape of molecules across the level $r$ will be obtained by replacing in (28) the integral

$$\int \cos\theta\, d\omega \qquad \text{.........(35)}$$

by the integral

$$\int \cos\theta\, [1 - f(r, \cos\theta)]\, d\omega \qquad \text{.........(36)},$$

in each case extended over the hemisphere. The integrand in (35) is however zero for $\theta \geqslant \theta_1$, and thus the value of (36) is

$$2\pi \int_0^{\theta_1} [1 - f(r, \cos\theta)] \cos\theta \sin\theta\, d\theta.$$

The value of (35) is $\pi$. Hence the corrected rate of escape will be obtained by multiplying (30) by the ratio of (36) to (35), *i.e.* by multiplying (30) by the function $\phi(r)$ given by†

$$\phi(r) = 2 \int_0^{\theta_1} [1 - f(r, \cos\theta)] \cos\theta \sin\theta\, d\theta \qquad \text{.........(37)}.$$

* Actually each molecule is moving along a hyperbola, but little error will be committed if we replace the portion of the hyperbola by the corresponding asymptote. In any case the curvature is small in the stretch of path that is important.

† If the height is so large that the molecular horizon is below the horizontal, $\theta_1$ must be replaced by $\frac{1}{2}\pi$ in these integrals.

For brevity put
$$\psi(r) = 2\rho r^2 \left(\frac{2\pi RT}{m}\right)^{\frac{1}{2}} e^{-q}(1+q) \quad \text{.............................(38).}$$

Then the rate of loss across the level $r$ in grams per second is $L(r)$, say, where
$$L(r) = \psi(r)\,\phi(r) \quad \text{.......................................(39).}$$

Now suppose $T \propto r^{-n}$. When $n > 1$, formulae (31) and (32) are valid generally, and we have
$$\psi(r) = 2\rho_0 r_0^2 \left(\frac{2\pi RT_0}{m}\right)^{\frac{1}{2}} \left(\frac{r}{r_0}\right)^{\frac{1}{2}n+2} (1+q)\, e^{\frac{q_0 - nq}{n-1}} \quad \text{....................(40).}$$

Since $q \propto r^{n-1}$, this expression tends steadily and rapidly to zero as $r \to \infty$, in virtue of the exponential factor.

When $n = 1$, $q$ is practically constant and equal to $q_0$ whilst $\rho$ is given by (§ 5)
$$\rho/\rho_0 = (r_0/r)^{q_0-1} \quad \text{..........................................(41).}$$

Hence when $n = 1$,
$$\psi(r) = 2\rho_0 r_0^2 \left(\frac{2\pi RT_0}{m}\right)^{\frac{1}{2}} \left(\frac{r_0}{r}\right)^{q_0 - \frac{3}{2}} (1+q_0)\, e^{-q_0} \quad \text{........................(42),}$$

which tends steadily and rapidly to zero as $r \to \infty$, since $q_0$ is large.

When $n < 1$, formulae (31) and (32) are not valid for large values of $r_0$ and we have to use the approximate formulae (§ 5)
$$\frac{\rho}{\rho_0} = \left(\frac{r_0}{r}\right)^{n+2} e^{-\frac{q_0 - q}{1-n}} = \left(\frac{r_0}{r}\right)^{n+2} e^{-\frac{q_0' - q'}{1-n}} \quad \text{..........................(43),}$$

where
$$q' = q - (2n+2), \quad q'/q_0' = (r_0/r)^{1-n} \quad \text{.............................(44).}$$

Hence for $n < 1$ we have
$$\psi(r) = 2\rho_0 r_0^2 \left(\frac{2\pi RT_0}{m}\right)^{\frac{1}{2}} \left(\frac{r_0}{r}\right)^{\frac{3}{2}n} (1+q)\, e^{-\frac{q_0 - nq}{1+n}} \quad \text{.......................(45).}$$

Here $q$ steadily decreases as $r \to \infty$, tending to the value $2n+2$. Hence in virtue of the exponential factor $\psi(r)$ steadily and rapidly decreases as $r \to \infty$ provided $n \neq 0$, but it only tends to zero in virtue of the factor $(r_0/r)^{\frac{3}{2}n}$. If $n = 0$, the exponential factor reduces to a constant, and $\psi(r)$ is nearly constant for a considerable range. Since, however, $q$ steadily decreases to the value 2, $\psi(r)$ steadily but slowly decreases to a small non-zero limit.

We now discuss the function $\phi(r)$. When $r = R$, the function $f(r, \cos\theta)$ is equal to unity for all values of $\theta$, and $\phi(r)$ is zero. When $r$ is very large, $f(r, \cos\theta)$ is nearly zero for all values of $\theta$, $\theta_1$ must be taken to be $\frac{1}{2}\pi$, and $\phi(r)$ is practically unity. Moreover for any given $\theta$ it is clear that $f(r, \cos\theta)$ steadily decreases as $r$ increases, and hence $\phi(r)$ steadily increases as $r$ increases.

Thus in general the rate of escape given by (39) is the product of two factors: the first factor $\psi(r)$ steadily decreases as $r \to \infty$, and tends to zero (or to a small positive value); the second factor $\phi(r)$ is zero when $r = R$, and steadily increases to the value unity as $r \to \infty$. *It follows that $L(r)$ has a maximum for some value of $r$ greater than $R$.* This maximum may be taken to give an upper limit to the rate of loss.

It should be noted that in all cases the behaviour of the function $\psi(r)$ during a considerable range of $r$ is practically that of the function $\rho^n$, which unless $n = 0$ is a decreasing function decreasing to zero. The function $\phi(r)$ is practically the available solid angle, an increasing function increasing from zero. These two statements taken together indicate the existence of a maximum. The anomalous case $n = 0$ arises because the function $\psi(r)$ is practically the product of the density $\rho$ into some function of the temperature and gravitational potential, and when $n = 0$ the latter function varies as $\rho^{-1}$ until $r$ becomes very large.

Even cruder arguments will show that $\psi(r)$ will in general ultimately tend to zero, and hence that $L(r)$ has a maximum. For if *all* the molecules at a given level had velocities which enabled them to escape, $\psi(r)$ would be proportional to the density times the area of the sphere, *i.e.* to $\rho r^2$. Actually, $\psi(r)$ must be a function which is always less than some multiple of $\rho r^2$. But, except when $n = 0$, $\rho r^2$ steadily decreases to zero as $r \to \infty$ (Section I). Hence $\psi(r)$ tends to zero. When $n = 0$, $\rho r^2$ tends to a non-zero limit, and the argument breaks down.

Jeans has pointed out that his formula—which is simply $\psi(r)$—for the rate of escape from an isothermal atmosphere (our formula (33) with $n = 0$) gives an always increasing rate of escape as $r$ increases*, and he has accounted for this as being due to its counting in, with increasing $r$, "satellite" molecules in free flight, which cannot really be supposed to be part of the genuine loss. Such molecules are of course counted in, from the nature of the method; but our more detailed analysis shows that the explanation is rather that the formula Jeans has used for the density $\rho$ (which does not take into account the effect of the mass of the atmosphere) gives the wrong behaviour for large values of $r$.

We are now in a position to regard the loss of molecules from the outer parts of an atmosphere from a more physical point of view. The analogy with evaporation is more close than is perhaps supposed, the main difference between an atmosphere and a liquid being that the former has no well-defined surface. But it is quite easy to define what may be called the "surface region" of an atmosphere. Any point taken below the height $R$ mentioned above may be regarded as being in the interior of the atmosphere, for there are no paths from it which avoid collisions. But above this height there will be a layer in which the sky near the zenith is clearing rapidly, and owing to the rapid fall-off of density it is obvious that the small circle $\theta = \theta_1$ separating the partially cleared sky from the lower uncleared sky will sink rapidly towards the horizon. This comparatively narrow layer in which $\theta_1$ changes quickly from 0 to a value near $\frac{1}{2}\pi$ may be regarded as the true surface of the atmosphere, and it is from this layer that evaporation (or loss of molecules) occurs. If the calculated maximum occurs in the region of rapid change of $\theta_1$ (where molecules are still comparatively abundant) we may adopt it with some confidence as giving the actual loss by diffusion.

Naturally the solution we thus obtain is only a crude approximation. The concept of the available solid angle is a loose one, and does not truly represent the state of affairs. For we postulate the complete equilibrium distribution of the gas to infinity, ignoring its molecular structure, and then use this steady state to determine a diffusion process, the occurrence of which is in contradiction with the existence of the steady state. However the complete steady state to infinity is only formally necessary, and if we confine the actual application of the formulae to what has been called above the "surface region" they be expected to give approximately correct results.

A further approximation could perhaps be obtained by taking into account the first collision (if any) of each of those molecules leaving a given level with a velocity exceeding the critical velocity, and calculating the proportion which still have escape-velocities after the collision. But it is questionable whether such a calculation would be worth the trouble. If further approximations were required the problem would have to be treated dynamically, as a genuine diffusion problem. Some suitable boundary condition would have to be formulated to represent the real

* In the text Jeans says that his formula (913) increases with $R$. This seems to be an oversight. Actually as it stands it decreases, but it increases when multiplied by $4\pi R^2$, for the area of the sphere.

state of affairs as well as possible—*e.g.* it might be supposed that the atmosphere was bounded by a given sphere $r = r_1$ and that molecules incident on the boundary with a velocity exceeding the critical velocity were supposed to be absorbed by the boundary.

§ 8. *Calculation of the solid angle subtended by the molecules above a given level.* It is necessary here to introduce a further approximation. In calculating the solid angle subtended by the molecules lying within any given solid angle we shall assume that each molecule subtends its own solid angle as though the other molecules were not present, so that overlapping is ignored. Overlapping will be negligibly small except when they appear to fill very nearly completely the solid angle considered. The circumstance that the molecules are in motion is obviously immaterial and is ignored in the calculation.

Let $\sigma$ be the diameter of a molecule (assumed spherical) and let $\nu(r)$ be the number of molecules per unit volume at the level $r$. Take the point $O$ at level $r$, and let $(x, \theta, \phi)$ be spherical polar co-ordinates with $O$ as origin and the normal through $O$ as axis. If $r'$ is the distance of any point $P(x, \theta, \phi)$ from the centre of the star, then

$$r'^2 = x^2 + r^2 + 2xr\cos\theta.$$

The effective solid angle subtended by a molecule, from the point of view of collisions with another equal molecule, is equal to that subtended by a sphere of *radius* $\sigma$. The solid angle subtended at $O$ by the molecules lying inside the element of volume $x^2 \sin\theta d\theta d\phi dx$ at $P$ is therefore

$$\frac{\nu(r')\pi\sigma^2}{x^2} x^2 \sin\theta d\theta d\phi dx,$$

and hence the solid angle subtended at $O$ by the molecules lying inside the solid angle

$$d\omega\ (= \sin\theta d\theta d\phi)$$

(assuming no overlapping) is

$$d\omega\pi\sigma^2 \int_0^\infty \nu(r')\, dx.$$

Hence

$$f(r, \cos\theta) = \pi\sigma^2 \int_0^\infty \nu(r')\, dx \quad \text{.......... (46).}$$

When $r$ does not exceed a certain value, $f(r, \cos\theta)$ will exceed unity even for $\theta = 0$; in this case overlapping occurs and the formula has no meaning. When $r$ is large, $f(r, \cos\theta)$ will be less than unity for a certain range of values of $\theta$; in this case the sky as viewed from $O$ will be partially clear down to a zenith distance $\theta_1$ given by

$$f(r, \cos\theta_1) = 1 \quad \text{.......... (47),}$$

provided this equation has a root*; and the sky will be completely covered in the range $\theta_1 \leqslant \theta \leqslant \frac{1}{2}\pi$. The total solid angle subtended by the molecules lying within the range $0 \leqslant \theta \leqslant \theta_1$ is $2\pi g(r)$, say, where

$$g(r) = \int_0^{\theta_1} f(r, \cos\theta) \sin\theta d\theta,$$

and the free solid angle is $\qquad 2\pi[1 - \cos\theta_1 - g(r)].$

However we are directly interested only in the derived function

$$\phi(r) = 2\int_0^{\theta_1} [1 - f(r, \cos\theta)] \cos\theta \sin\theta d\theta$$

$$= 1 - \cos^2\theta_1 - 2\phi_1(r) \quad \text{.......... (48),}$$

* If $f(r, \cos\theta) < 1$ for all values of $\theta$ lying between 0 and $\frac{1}{2}\pi$, $\theta_1$ is defined as being equal to $\frac{1}{2}\pi$.

say, where $$\phi_1(r) = \int_0^{\theta_1} f(r, \cos\theta)\cos\theta\sin\theta d\theta \quad \text{.........(49)}.$$

We shall now obtain asymptotic approximations for this for the different cases that arise.

*Case I.* $n > 1$. We have $\nu = \rho/m$, and $\rho$ is given by (31) in terms of $\rho_0$ and $q_0$ at any arbitrary reference height $r_0$. Take this reference height to be that of the point $O$ we are considering, substitute for $\rho$ in (46) and then omit the suffix 0 wherever it occurs. We find

$$f(r, \cos\theta) = \frac{\pi\sigma^2\rho}{mr^n}\int_0^\infty (x^2 + r^2 + 2xr\cos\theta)^{\frac{1}{2}n} e^{-\frac{q}{n-1}\left[\left(\frac{x^2+r^2+2xr\cos\theta}{r^2}\right)^{\frac{1}{2}(n-1)} - 1\right]} dx \quad \text{...(50)},$$

where $r$, $\rho$ and $q$ are the values at $O$. Now substitute $x = r\xi$, and put $\cos\theta = \mu$. Then

$$f(r, \mu) = \frac{\pi\sigma^2\rho r}{m}\int_0^\infty (1 + 2\mu\xi + \xi^2)^{\frac{1}{2}n} e^{-\frac{q}{n-1}[(1+2\mu\xi+\xi^2)^{\frac{1}{2}(n-1)} - 1]} d\xi \quad \text{.........(51)}.$$

In this put $$1 + 2\mu\xi + \xi^2 = (1 + y/\alpha)^{2/(n-1)},$$
where $$\alpha = q/(n-1).$$

We find then $$f(r, \mu) = \frac{\pi\sigma^2\rho r}{mq}\int_0^\infty \frac{(1 + y/\alpha)^{3/(n-1)} e^{-y}}{[(1 + y/\alpha)^{2/(n-1)} - 1 + \mu^2]^{\frac{1}{2}}} dy \quad \text{.........(52)}.$$

It will be seen in Section III that in practice $q$ is large compared with unity in the boundary region, and therefore, since in practice $n \leqslant 2$ (§ 1), $\alpha$ is also large compared with unity. If the integrand in (52) is expanded in powers of $\alpha^{-1}$ and integrated term by term, the result will be an asymptotic expansion valid when $q$ is large. The first term of this expansion is sufficient for our purpose. Provided $\mu \neq 0$, we find in fact for $q$ large

$$f(r, \mu) \sim \frac{\pi\sigma^2\rho r}{\mu mq}\left[1 + \frac{3\mu^2 - 1}{\mu^2 q} + \dots\right] \quad \text{.........(53)},$$

and hence so long as $\mu^2 q$ is not too small we may take

$$f(r, \mu) = \frac{\pi\sigma^2\rho r}{mq} \cdot \frac{1}{\mu} \quad \text{.........(54)}.$$

Notice that the approximate form (54) and even the second term in (53) are independent of $n$; the complete expression of $f(r, \mu)$ however would obviously involve $n$.

Using expression (54) in (47) we have for the position of the molecular horizon

$$\cos\theta_1 = \mu_1 = \frac{\pi\sigma^2\rho r}{mq} \quad \text{.........(55)}.$$

The expression on the right-hand side of (55) must not exceed unity for $\theta_1$ to be real. Thus the foregoing calculations are valid provided $r$ is so large ($\rho$ so small) that

$$\frac{\pi\sigma^2\rho r}{mq} \leqslant 1 \quad \text{.........(56)},$$

and provided also that $r$ is still sufficiently small for

$$\mu_1^2 q \quad \text{or} \quad \frac{1}{q}\left(\frac{\pi\sigma^2\rho r}{m}\right)^2 \quad \text{.........(57)}$$

to be large compared with unity. It is convenient to express $f(r, \mu)$ in the approximate form

$$f(r, \mu) = \mu_1/\mu \quad \text{.........(58)}.$$

Since $\rho$ varies rapidly with the height (more rapidly than $q$), therefore *there is a certain range of height in which $\mu_1$ varies comparatively rapidly from unity to a value near zero*, as already anticipated in § 7.

If $r$ is so large that it is necessary to consider small values of $\mu$, condition (57) may be violated and we cannot proceed to approximation (53). When $\mu^2 q$ is not large compared with unity the leading terms in (52) give

$$f(r, \mu) \sim \frac{\pi\sigma^2\rho r}{mq} \int_0^\infty \left(\frac{2y}{q} + \mu^2\right)^{-\frac{1}{2}} e^{-y} dy \quad \text{.........................(59)},$$

which if $\mu^2 q$ is small compared with unity may be approximated to in the form

$$f(r, \mu) \sim \frac{\pi\sigma^2\rho r}{m} \left(\frac{\pi}{2q}\right)^{\frac{1}{2}} \left[1 - 2\left(\frac{\frac{1}{2}q\mu^2}{\pi}\right)^{\frac{1}{2}}\right] \quad \text{.........................(60)}.$$

It follows from this that $f(r, 0) < 1$ when $r$ is so large that

$$\frac{\pi\sigma^2\rho r}{mq} < \left(\frac{2}{\pi q}\right)^{\frac{1}{2}} \quad \text{.........................................(61)},$$

which is therefore the condition that the molecular horizon shall be below the horizontal.

To find $\phi_1(r)$ we have, using (52),

$$\phi_1(r) = \int_{\mu_1}^1 f(r, \mu)\, \mu d\mu = \frac{\pi\sigma^2\rho r}{mq} \int_{\mu_1}^1 \mu d\mu \int_0^\infty \frac{(1 + y/\alpha)^{3/(n-1)} e^{-y}}{[(1 + y/\alpha)^{2/(n-1)} - 1 + \mu^2]^{\frac{1}{2}}} dy \quad \text{......(62)}.$$

Inverting the order of integration and integrating with respect to $\mu$ we find

$$\phi_1(r) = \frac{\pi\sigma^2\rho r}{mq} \int_0^\infty [(1 + y/\alpha)^{1/(n-1)} - \{(1 + y/\alpha)^{2/(n-1)} - 1 + \mu_1^2\}^{\frac{1}{2}}] (1 + y/\alpha)^{3/(n-1)} e^{-y} dy \quad \text{...(63)}.$$

The asymptotic expansion of this valid for large values of $\mu_1^2 q$ is found to be

$$\phi_1(r) \sim \frac{\pi\sigma^2\rho r}{mq} (1 - \mu_1) \left[1 + \frac{3\mu_1 - 1}{\mu_1 q} + \ldots\right] \quad \text{.......................(64)}.$$

Hence we have approximately

$$\phi_1(r) = \frac{\pi\sigma^2\rho r}{mq} (1 - \mu_1)$$
$$= \mu_1(1 - \mu_1) \quad \text{..........................................(65)}.$$

The same expression would have been obtained if we had used the approximate formula for $f(r, \mu)$ in calculating $\phi_1(r)$, thus

$$\phi_1(r) = \int_{\mu_1}^1 f(r, \mu)\, \mu d\mu = \int_{\mu_1}^1 \frac{\mu_1}{\mu} \mu d\mu = \mu_1(1 - \mu_1).$$

If $\mu_1^2 q$ is not large compared with unity, the leading term in the expansion of (63) gives

$$\phi_1(r) \sim \frac{\pi\sigma^2\rho r}{mq} \int_0^\infty \left[1 - \left(\mu_1^2 + \frac{2y}{q}\right)^{\frac{1}{2}}\right] e^{-y} dy \quad \text{.....................(66)},$$

which again if $\mu_1^2 q$ is small compared with unity gives

$$\phi_1(r) \sim \frac{\pi\sigma^2\rho r}{mq} \left[1 - \left(\frac{\pi}{2q}\right)^{\frac{1}{2}}\right] \quad \text{...............................(67)}.$$

We shall find that the most important case is when $\mu_1$ is not nearly zero. Summarising for this case, we have the approximate formulae

$$\left.\begin{aligned} \cos\theta_1 &= \mu_1 = \frac{\pi\sigma^2\rho r}{mq}, \\ f(r, \mu) &= \frac{\pi\sigma^2\rho r}{mq\mu} = \frac{\mu_1}{\mu}, \\ \phi_1(r) &= \mu_1(1 - \mu_1), \\ \phi(r) &= 1 - \mu_1^2 - 2\phi_1(r) = (1 - \mu_1)^2 \end{aligned}\right\} \quad \text{.........................(68)}.$$

*Case II.* $n=1$. Either by substituting the appropriate formula for $\rho$ from § 5, or by passing to the limit $(n \to 1)$ in (50) or (51) we find for $n=1$

$$f(r, \mu) = \frac{\pi\sigma^2\rho r}{m}\int_0^\infty (1+2\mu\xi+\xi^2)^{-\frac{1}{2}(q-1)}\,d\xi \quad \text{.......................(69)},$$

which converges since $q$ is large. Substituting

$$\tfrac{1}{2}(q-3)\log(1+2\mu\xi+\xi^2) = y,$$

this becomes

$$f(r, \mu) = \frac{\pi\sigma^2\rho r}{m(q-3)}\int_0^\infty \frac{e^{-y}\,dy}{(e^{2y/(q-3)}-1+\mu^2)^{\frac{1}{2}}} \quad \text{.......................(70)}.$$

The leading terms in the asymptotic expansion of this, for $\mu^2 q$ large or small, are the same as for the case $n>1$ just discussed. The same applies to the expression which can be obtained for $\phi_1(r)$. Hence formulae (68) are valid when $n=1$.

*Case III.* $n<1$. In this case the integrand in (50) does not tend to zero, and the integrals already obtained for $f(r, \mu)$ and $\phi_1(r)$ diverge. This would appear to indicate that all elementary solid angles are completely filled with molecules. But this conclusion would be erroneous; it arises from the fact that the expression for the density valid when $n>1$ is not valid for $n<1$ for large values of $r$. We must accordingly use in (46) the approximation for $\rho$ which gives the correct behaviour for $r$ large, as obtained in Section I. The approximation in question has already been quoted in Section II, in equations (43) and (44). Employing these in (46), then omitting the suffix 0 and substituting

$$x = r\xi, \quad \alpha' = q'/(1-n),$$

we find

$$f(r, \mu) = \frac{\pi\sigma^2\rho r}{m}\int_0^\infty (1+2\mu\xi+\xi^2)^{-\frac{1}{2}(n+2)}\,e^{-\alpha'[1-(1+2\mu\xi+\xi^2)^{-\frac{1}{2}(1-n)}]}\,d\xi \quad \text{......(51')}.$$

This integral converges. The exponential factor tends to a non-zero limit as $\xi \to \infty$, but convergence occurs owing to the first factor in the integrand, since $n \geqslant 0$. Put

$$1+2\mu\xi+\xi^2 = (1-y/\alpha')^{-2/(1-n)}.$$

Then

$$f(r, \mu) = \frac{\pi\sigma^2\rho r}{mq'}\int_0^{\alpha'} \frac{(1-y/\alpha')^{2n/(1-n)}\,e^{-y}\,dy}{[1-(1-\mu^2)(1-y/\alpha')^{2/(1-n)}]^{\frac{1}{2}}} \quad \text{.................. (52')}.$$

This should be contrasted with (52). When $\mu^2 q'$ is large, $f(r, \mu)$ has the asymptotic expansion

$$f(r, \mu) \sim \frac{\pi\sigma^2\rho r}{\mu m q'}\left[1 - \frac{(2n-1)\mu^2+1}{\mu^2 q'} + \dots\right] \quad \text{.......................(53')}.$$

Inserting $q' = q-(2n+2)$ and expanding in powers of $q^{-1}$, (53') becomes

$$f(r, \mu) \sim \frac{\pi\sigma^2\rho r}{\mu m q}\left[1 + \frac{3\mu^2-1}{\mu^2 q} + \dots\right],$$

of which the first two terms are identical with those of the corresponding expansion for $n>1$, namely (53).

When $\mu^2 q'$ is not large compared with unity the leading term in the expansion of (51') gives

$$f(r, \mu) \sim \frac{\pi\sigma^2\rho r}{mq}\int_0^{\alpha'}\left[\mu^2 + \frac{(1-\mu^2)\,y}{\frac{1}{2}q'}\right]^{-\frac{1}{2}} e^{-y}\,dy \quad \text{.......................(59')},$$

and when $\mu^2 q'$ is small compared with unity this has the approximate expansion (60), the same as when $n>1$.

To find $\phi_1(r)$, using the expression of $f(r, \mu)$ given by (51′), inverting the order of integration and integrating with respect to $\mu$ we find

$$\phi_1(r) = \int_{\mu_1}^{1} f(r, \mu)\, \mu d\mu = \frac{\pi\sigma^2 \rho r}{mq'} \int_0^{\alpha'} \frac{1 - [1 - (1-\mu_1^2)(1 - y/\alpha')^{2/(1-n)}]^{\frac{1}{2}}}{(1 - y/\alpha')^2} e^{-y} dy$$

$$= \frac{\pi\sigma^2 \rho r}{mq'} (1 - \mu_1^2) \int_0^{\alpha'} \frac{(1 - y/\alpha')^{2n/(1-n)}}{1 + [1 - (1-\mu_1^2)(1 - y/\alpha')^{2/(1-n)}]^{\frac{1}{2}}} e^{-y} dy \quad \text{.........(63′)}.$$

This converges. It should be contrasted with (63). The asymptotic expansion of this valid for large values of $\mu_1^2 q'$ is found to be

$$\phi_1(r) \sim \frac{\pi\sigma^2 \rho r}{mq'} (1 - \mu_1) \left[1 - \frac{(2n-1)\mu_1 + 1}{\mu_1 q'} + \ldots\right] \quad \text{..................(64′)}.$$

Inserting $q' = q - (2n+2)$ and expanding in powers of $q^{-1}$, (64′) becomes

$$\phi_1(r) \sim \frac{\pi\sigma^2 \rho r}{mq} (1 - \mu_1) \left[1 + \frac{3\mu_1 - 1}{\mu_1 q} + \ldots\right],$$

of which the first two terms are identical with those of the corresponding expansion for $n > 1$, namely (64).

When $\mu_1^2 q'$ is not large compared with unity the leading term in the expansion of (63′) gives

$$\phi_1(r) \sim \frac{\pi\sigma^2 \rho r}{mq'} \int_0^{\alpha'} \left[1 - \left(\mu_1^2 + \frac{1-\mu_1^2}{\frac{1}{2}q'} y\right)^{\frac{1}{2}}\right] e^{-y} dy \quad \text{..................(66′)},$$

which again if $\mu_1^2 q'$ is small compared with unity gives

$$\phi_1(r) \sim \frac{\pi\sigma^2 \rho r}{mq'} \left[1 - \left(\frac{\pi}{2q'}\right)^{\frac{1}{2}}\right] \quad \text{..............................(67′)}.$$

If in this we put $q' = q - (2n+2)$, the resulting expansion is identical with (67) as far as the power of $q^{-1}$ concerned.

We thus see that the approximate formulae (68) hold for $0 \leqslant n \leqslant 1$ as well as for $n > 1$. The fact that (68) do not involve $n$ explicitly might have suggested that this would be the case, but the method of derivation of (68) for $n > 1$ does not hold for $n < 1$, and a separate investigation was necessary.

§ 9. *The maximum of the formal expression for the rate of escape.* We can now investigate in detail the formal expression for the rate of escape as given by (39). Let $r_0$ as usual be some convenient reference level. Assume in the first instance that the level $r$ is such that conditions (56) and (57) are satisfied.

*Case I.* $n > 1$. Using (40) and (68) we have

$$L(r) = 2\rho_0 r_0^2 \left(\frac{2\pi RT_0}{m}\right)^{\frac{1}{2}} \left(\frac{r}{r_0}\right)^{\frac{1}{2}n+2} e^{\frac{q_0 - nq}{n-1}} (1+q)(1-\mu_1)^2 \quad \text{..............(71)}.$$

Now
$$\mu_1 = \frac{\pi\sigma^2 \rho r}{mq} = \frac{\pi\sigma^2 \rho_0 r_0}{mq_0} \left(\frac{r}{r_0}\right)^2 e^{\frac{q_0 - q}{n-1}} \quad \text{..............................(72)},$$

on substituting for $\rho$ and $q$ from (31). Eliminating the factor $e^{-q/(n-1)}$ between (71) and (72) we have

$$L(r) = 2\rho_0 r_0^2 \left(\frac{2\pi RT_0}{m}\right)^{\frac{1}{2}} \left(\frac{mq_0}{\pi\sigma^2 \rho_0 r_0}\right)^n \left(\frac{r_0}{r}\right)^{\frac{3}{2}n-2} (1+q)\, e^{-q_0} \mu_1^n (1-\mu_1)^2 \quad \text{........(73)}.$$

In this expression the factor $(r_0/r)^{\frac{3}{2}n-2}(1+q)$ varies practically as $(r_0/r)^{\frac{1}{2}n-1}$, and so varies slowly compared with $\mu_1$. Hence the variation of (73) is dominated by the factor

$$\zeta(\mu_1) = \mu_1^n (1-\mu_1)^2 \quad \text{........................................(74)}.$$

This is sensitive to the value of $r$, since $\mu_1$ over any limited range is approximately proportional to $\rho$. It vanishes for $\mu_1 = 1$ and $\mu_1 = 0$. Differentiating for the maximum lying in the range $0 < \mu_1 < 1$, we find

$$(\mu_1)_{\max} = \frac{n}{n+2} \quad \text{.........................................(75).}$$

A few numerical values (not restricted to $n > 1$ in view of later work) are given in the following table.

TABLE I.

| $n$ | $(\mu_1)_{\max}$ | $(\theta_1)_{\max}$ | $\zeta_{\max}$ |
|---|---|---|---|
| 0 | 0 | 90° | 1 |
| $\frac{1}{10}$ | $\frac{1}{21}$ | 87° 17′ | 0·668 |
| $\frac{1}{4}$ | $\frac{1}{9}$ | 83° 37′ | 0·457 |
| $\frac{1}{2}$ | $\frac{1}{5}$ | 78° 25′ | 0·286 |
| $\frac{2}{3}$ | $\frac{1}{4}$ | 75° 31′ | 0·223 |
| 1 | $\frac{1}{3}$ | 70° 32′ | 0·148 |
| $\frac{3}{2}$ | $\frac{3}{7}$ | 64° 37′ | 0·0917 |
| 2 | $\frac{1}{2}$ | 60° | 0·0625 |

In the range $1 < n \leq 2$ $(\mu_1)_{\max}$ lies between $\frac{1}{3}$ and $\frac{1}{2}$, and therefore $\mu_1{}^2 q$ is sufficiently large for the approximations we have used to be valid.

Thus for $n > 1$ the maximum of the formal expression for the rate of escape occurs for a value of $\theta_1$ in the neighbourhood of 60°—70°. Such values of $\theta_1$ correspond to the layer in which the sky of an ascending observer is clearing rapidly—the surface region of the atmosphere. Thus the formal maximum occurs in the region from which evaporation (if any) must be chiefly proceeding. The maximum may therefore be taken as giving the true magnitude of the rate of escape.

The density at the height of the formal maximum may be found by substituting the value of $(\mu_1)_{\max}$ in (72). Omitting the suffix *max* for brevity we have

$$\rho = \frac{mq\mu_1}{\pi\sigma^2 r} \quad \text{.........................................(76).}$$

In this $r_0$ and $q_0$ may be substituted for $r$ and $q$ with an inappreciable error, and we have finally

$$\rho = \frac{mq_0\mu_1}{\pi\sigma^2 r_0} \quad \text{.........................................(77),}$$

a formula which evaluates $\rho$ at the escape level in terms of known quantities.

According to the theory of gases, the mean free path $l$ at a given density $\rho$ is given by the formula

$$l = \frac{m}{\sqrt{2}\pi\sigma^2\rho} \quad \text{.........................................(78).}$$

Dividing (77) by (78) we find

$$\frac{l}{r_0} = \frac{1}{\sqrt{2}\mu_1 q_0} \quad \text{.........................................(79).}$$

The simplicity of this formula is worthy of note. The "mean free path" thus found is not, of course, the actual mean free path at the height concerned. It is the mean free path which would occur in a large amount of gas of uniform density equal to the density given by (77). When the density changes appreciably in a distance comparable with the value of $l$ given by (78), the latter formula no longer applies, the mean free path being in fact different in different directions. However (79) gives the equivalent mean free path at the height at which the molecular horizon is given by $\mu_1$, and the use of the equivalent mean free path is perhaps the simplest way of visualising the order of magnitude of the density there.

To find the height at which the formal maximum occurs we must use (72). Taking logarithms,

$$\frac{q_0}{n-1}\left(1-\frac{r^{n-1}}{r_0^{\,n-1}}\right)=\log\frac{mq_0\mu_1}{\pi\sigma^2\rho_0 r_0}+2\log\frac{r_0}{r}.$$

Putting $r/r_0 = 1 + z$, where $z$ is small, we find

$$(q_0-2)\,z=\log\frac{\pi\sigma^2\rho_0 r_0}{mq_0\mu_1},$$

or, since $q_0$ is large compared with 2,

$$\frac{r-r_0}{r_0}=z=\frac{1}{q_0}\log\frac{\pi\sigma^2\rho_0 r_0}{mq_0\mu_1} \qquad\text{.................................(80).}$$

From numerical values given in Section III it will appear that, owing to the largeness of $q_0$, $z$ is in fact small compared with unity, as assumed.

In expression (73) for the magnitude of the rate of loss we can now put $r = r_0$, $q = q_0$, approximately. We have then

$$L(r)_{\max}=2\rho_0 r_0^{\,2}\left(\frac{2\pi RT_0}{m}\right)^{\frac{1}{2}}\left(\frac{mq_0}{\pi\sigma^2\rho_0 r_0}\right)^{n} q_0 e^{-q_0}\,\zeta_{\max} \qquad\text{....................(81),}$$

where

$$\zeta_{\max}=\frac{4n^n}{(n+2)^{n+2}}.$$

This gives the rate of loss in terms of quantities depending only on the reference height.

*Case II.* $n < 1$. We shall not pause to discuss case $n = 1$ separately. When $n < 1$, using (45) and (68), we have

$$L(r)=2\rho_0 r_0^{\,2}\left(\frac{2\pi RT_0}{m}\right)^{\frac{1}{2}}\left(\frac{r_0}{r}\right)^{\frac{3}{2}n} e^{-\frac{q_0-nq}{1-n}}(1+q)(1-\mu_1)^2 \qquad\text{.............(71'),}$$

where (substituting for $\rho$ and $q$ from (43) in (68))

$$\mu_1=\frac{\pi\sigma^2\rho r}{mq}=\frac{\pi\sigma^2\rho_0 r_0}{mq_0}\left(\frac{r_0}{r}\right)^{n+1}\left(\frac{q_0}{q}\right)e^{\frac{q_0-q}{1-n}} \qquad\text{.......................(72').}$$

Eliminating the factor $e^{q/(1-n)}$ between (71′) and (72′) we have

$$L(r)=2\rho_0 r_0^{\,2}\left(\frac{2\pi RT_0}{m}\right)^{\frac{1}{2}}\left(\frac{mq_0}{\pi\sigma^2\rho_0 r_0}\right)^{n}\left(\frac{r}{r_0}\right)^{n^2-\frac{1}{2}n}\left(\frac{q}{q_0}\right)^{n}(1+q)\,e^{-q}\mu_1^{\,n}(1-\mu_1)^2 \;\text{......(73').}$$

Provided $r/r_0$ is not large, $(q/q_0)^n$ is approximately $(q'/q_0')^n$ which is equal to $(r_0/r)^{n-n^2}$. Thus the variation of (73′) is dominated by the same factor as when $n > 1$, namely the factor $\zeta(\mu_1) = \mu_1^{\,n}(1-\mu_1)^2$, whose maximum for various values of $n$ has been given above. But when $n$ is nearly zero, the value of $(\mu_1)_{\max}$ as given by (75) is very small, hence $\mu_1^{\,2}q$ in the region of the maximum is not necessarily large, and so the approximation used for $\phi(r)$ is no longer valid. A fresh investigation is therefore required when $n$ is zero or nearly zero. When $n < 1$ there are thus two sub-cases to consider.

*Case II a. n not nearly zero.* Although (71′) is not quite identical with (71), the ratio is practically only a power of $r/r_0$. The height at which the maximum occurs is such that $r/r_0$ differs but little from unity, and the discussion of the density and equivalent mean free path at the maximum, of the height of the maximum and of the actual value of the loss is the same as for $n > 1$. Formulae (76)...(81) apply as they stand.

*Case II b. n zero or nearly zero.* Suppose that $r$ is so large that $\mu_1^2 q$ is small or zero. Then we must use the approximation (67′), giving since $\mu_1$ is small or zero

$$\phi(r) = 1 - 2\phi_1(r) = 1 - \frac{2\pi\sigma^2\rho r}{mq'},$$

and the formal expression for the rate of escape is

$$L(r) = 2\rho_0 r_0^2 \left(\frac{2\pi RT_0}{m}\right)^{\frac{1}{2}} \left(\frac{r_0}{r}\right)^{\frac{3}{2}n} e^{-\frac{q_0 - nq}{1-n}} (1+q)\left(1 - \frac{2\pi\sigma^2\rho r}{mq'}\right) \quad \text{..........(82).}$$

Making the proper substitutions for $\rho$ and $q$, putting $r/r_0 = 1 + z$ and differentiating logarithmically, we find the maximum must occur when

$$-nq_0'(1+z)^{n-2} - \frac{\frac{3}{2}n}{1+z} - \frac{q_0'(1-n)(1+z)^{n-2}}{2n+3+q'} + \frac{2\pi\sigma^2\rho r}{mq'} \cdot \frac{q_0'(1+z)^{n-2} + 2n(1+z)^{-1}}{1 - 2\pi\sigma^2\rho r/mq'} = 0.$$

Neglecting $z$ wherever it occurs explicitly in this equation, and neglecting also $n/q_0'$ compared with unity, we find that the maximum must occur for

$$\frac{\pi\sigma^2\rho r}{mq'} = \tfrac{1}{2}\left(n + \frac{1+\frac{1}{2}n}{q_0'}\right)\bigg/\left(1 + n + \frac{1+\frac{1}{2}n}{q_0'}\right) \quad \text{..........(83).}$$

Since $L(r)$ is only given by (82) provided $\pi\sigma^2\rho r/mq'$ is small compared with unity, this formula for the maximum is only valid if $\frac{1}{2}\left(n + \frac{1}{q_0'}\right)$ is small compared with unity; as we have assumed $n$ to be zero or nearly zero, this condition is satisfied. We may therefore write (83) in the form

$$\frac{\pi\sigma^2\rho r}{mq'} = \tfrac{1}{2}\left(n + \frac{1}{q_0}\right) \quad \text{..........(84),}$$

approximately. It should be observed that when $n$ though small compared with unity is large compared with $1/q_0$, formula (84) agrees with (75), both reducing to $\pi\sigma^2\rho r/mq = \frac{1}{2}n$.

From (84) the equivalent mean free path at the maximum is found to be given by

$$\frac{l}{r_0} = \frac{\sqrt{2}}{nq_0 + 1} \quad \text{..........(85).}$$

To find the height of the maximum we have

$$\frac{\rho r}{q'} = \frac{\rho_0 r_0}{q_0'}\left(\frac{r_0}{r}\right)^{2n} e^{-\frac{q_0' - q'}{1-n}},$$

whence inserting in (84) we find approximately

$$\frac{r - r_0}{r_0} = z = \frac{1}{q_0}\log\frac{\pi\sigma^2\rho_0 r_0}{mq_0} + \frac{1}{q_0}\log\frac{2q_0}{nq_0+1} \quad \text{..........(86).}$$

(The numerical results of Section III will justify our assumption that even in this case $z$ is small.)

The magnitude of $L(r)$ at the formal maximum is found to be approximately

$$L(r) = 2\rho_0 r_0^2 \left(\frac{2\pi RT_0}{m}\right)^{\frac{1}{2}} \left(\frac{mq_0}{\pi\sigma^2\rho_0 r_0}\right)^n q_0 e^{-q_0}\left[\tfrac{1}{2}\left(n + \frac{1}{q_0}\right)\right]^n \quad \text{..........(87).}$$

Comparison of (86) with (80) will show that when $n$ is very small or zero the height at which the formal maximum occurs is much greater than when $n$ has larger values. Likewise comparison of (84) with (77) and of (85) with (79) will show that the density and equivalent mean free path at the maximum are respectively much smaller and much longer.

However, when $n$ is very small or zero the maximum is not a sharp one; there is a considerable range of $r$ in which the formal expression for the rate of loss is nearly constant. It is questionable, therefore, whether any advantage is gained by regarding the height at which the formal maximum occurs as the height from which evaporation is mainly proceeding. Almost the same value for the rate of loss will be obtained whether we evaluate $L(r)$ at its maximum or whether we evaluate it in the region in which $\mu_1$ changes rapidly from unity to zero. Physically it is the latter region to which the term "surface region" is best applicable, and evaporation from this region is practically the whole of the evaporation. If we put, say, $\mu_1 = \frac{1}{10}$, ($\theta_1 = 84°$) corresponding strictly to $n = \frac{2}{9}$, the formulae of Case I for $\rho$ and $l$ will represent $n = 0$ sufficiently accurately. The factor $\zeta_{\max}$ in (81) has the same limit (unity) as $n \to 0$ as the corresponding factor in (87).

§ 10. *Discussion of the formula for the rate of loss.* Formulae (81) and (87) show an explicit dependence of the rate of loss on the size of the molecules except when $n = 0$, smaller molecules in general escaping more rapidly. This would perhaps be expected, but it is difficult to see why the loss should be independent of $\sigma$ when the atmosphere is isothermal.

The factor $(mq_0/\pi\sigma^2\rho_0 r_0)^n$ is equal to $(\rho/\mu_1\rho_0)^n$, and unless $n = 0$ is very small, since $\rho/\rho_1$ will be small, $\rho$ being the density at the escape level. Thus if $n$ is but a little different from zero, the rate of loss is much smaller than in the isothermal case.

The rate of loss is proportional to $\rho_0^{1-n}$. Thus if $n = 0$, the rate is directly proportional to $\rho_0$, in agreement with Jeans. If $n = 1$, the rate is independent of $\rho_0$. The time for an amount of gas to be lost equal to the whole amount of atmosphere above $r_0$ is proportional to $\rho_0^n$. However, strictly speaking the removal of each molecule causes a slight re-adjustment of the whole atmosphere, with consequent diminution of $\rho_0$*, and thus in calculating the time for a whole outer atmosphere to stream away $\rho_0$ would have to be taken as a function of the time. This function could probably be determined without trouble. When $n = 1$ this question does not arise, the rate of loss being independent of $\rho_0$.

§ 11. *Summary of the mathematical results of Section II.* As seen from a point at a height $r$ in the atmosphere, where the density is $\rho$, the zenith distance $\theta_1$ of the molecular horizon is given approximately by

$$\cos\theta_1 = \mu_1 = \frac{\pi\sigma^2\rho r}{mq} \quad \text{.....(88)},$$

provided that $\rho$ is so small that $\mu_1 < 1$ but not so small that $\mu_1$ is nearly 0. Inside the zenith distance $\theta_1$ the sky is everywhere partially free of molecules. In this formula $\sigma$ and $m$ are the diameter and mass of a molecule, and $q$ is given by

$$q = mV/RT,$$

where $V$ is the gravitational potential and $T$ is the temperature, which varies as $r^{-n}$.

To take account of the free solid angle available for escaping molecules, the rate of escape calculated as though the whole outer hemisphere were free must be multiplied by the factor $(1-\mu_1)^2$. The formal expression thus obtained for the rate of escape has a maximum at a density $\rho$, height $r$, zenith distance of molecular horizon $\theta_1$ and equivalent mean free path $l$ given by

$$\cos\theta_1 = \mu_1 = \frac{n}{n+2} \quad \text{.....(89)},$$

$$\rho = \frac{mq_0\mu_1}{\pi\sigma^2 r_0} \quad \text{.....(90)},$$

$$\frac{r}{r_0} = 1 + \frac{1}{q_0}\log\frac{\pi\sigma^2\rho_0 r_0}{mq_0\mu_1} = 1 + \frac{1}{q_0}\log\frac{\rho_0}{\rho} \quad \text{.....(91)},$$

$$\frac{l}{r_0} = \frac{1}{\sqrt{2q_0\mu_1}} \quad \text{.....(92)},$$

* I owe this remark to Dr J. E. Jones, of Trinity College, Cambridge.

and the magnitude of the rate of escape at the maximum, in grams per second from the whole surface of the atmosphere, is given by

$$L(r) = 2\rho_1 r_0^2 \left(\frac{2\pi R T_0}{m}\right)^{\frac{1}{2}} \left(\frac{mq_0}{\pi\sigma^2\rho_0 r_0}\right)^n q_0 e^{-q} \frac{4n^n}{(n+2)^{n+2}} \quad \text{..............(93)}.$$

The suffix 0 refers to the height selected as reference height.

The formulae of the last paragraph are not valid if $n$ is very small or zero. In this case the maximum occurs at a height at which the molecular horizon is near or below the horizontal. The value of $\mu_1$ at the maximum is no longer given by (89), but the correct formulae for $\rho$, $r$ and $l$ at the maximum are obtained by replacing $\mu_1$ in (90), (91) and (92) by the quantity

$$\tfrac{1}{2}\left(n + \frac{1}{q_0}\right).$$

The value of $L(r)$ at the maximum is obtained by replacing the last factor in (93) by the factor

$$\left[\tfrac{1}{2}\left(n + \frac{1}{q_0}\right)\right]^n.$$

### III. *Applications.*

§ 12. *The gravitational potential at the surface of a star.* The value of $q_0$, on which the order of magnitude of the loss depends, is given by

$$q_0 = GmM/r_0 RT_0 = mV_0/RT_0 \quad \text{..............................(94)}.$$

To search for stars from the surfaces of which the loss by diffusion might be appreciable is to search for stars having low values of $q_0$. It is plain that if we compare a giant and a dwarf of the same mass and temperature, the value of $q_0$ will be much smaller for the giant than for the dwarf (*e.g.* some eight times smaller in the case of the sun and its corresponding giant). We can therefore confine our immediate attention to the giants. If $T_1$ is the effective temperature, $k$ the mean coefficient of absorption in the interior, we have on Eddington's theory of the radiative equilibrium of a giant star

$$T_1^4 = \frac{4GM}{r_0^2}\,\frac{1-\beta}{ak} \quad \text{..........................................(95)},$$

where $1-\beta$ is the ratio of the radiation pressure to the combined hydrostatic and radiation pressure. It is connected with the mass by the relation

$$1-\beta = 0{\cdot}0026\,(M/\odot)^2 m'^4 \beta^4 \quad \text{..................................(96)},$$

where $m'$ denotes the mean molecular weight of the gas in the interior of the star and $\odot$ denotes the mass of the sun; $a$ is such that $aT^4$ is the energy-density of black radiation at temperature $T$. From (94) and (95) we have

$$V_0 = \frac{GM}{r_0} = \tfrac{1}{2}T_1^2\left(\frac{akGM}{1-\beta}\right)^{\frac{1}{2}} \quad \text{..................................(97)},$$

$$q_0 = \tfrac{1}{2}\frac{m}{R}\,\frac{T_1^2}{T_0}\left(\frac{akGM}{1-\beta}\right)^{\frac{1}{2}} \quad \text{................................(98)}.$$

For grey material $T_0 = 2^{-\frac{1}{4}}T_1$; and whatever the optical properties it has been shown (under certain conditions which are probably satisfied) that*

$$T_1 > T_0 > \tfrac{1}{2}T_1.$$

* Milne, *Monthly Notices*, 82, 368, 1922.

For simplicity we shall take $T_0 = 2^{-\frac{1}{4}} T_1 = 0\cdot84 T_1$. We have then

$$q_0 = \frac{2^{-\frac{3}{4}} m T_1}{R} \left(\frac{akGM}{1-\beta}\right)^{\frac{1}{2}} \qquad \text{(99)}.$$

Hence for giant stars of given mass, $q_0 \propto T_1$; and therefore to obtain small values of $q_0$ we must look amongst the stars of lowest temperature.

It is next necessary to examine the variation of $q_0$ or ($V_0$) with $M$, for stars of given temperature. When $M$ is large, $(1-\beta)$ tends to unity, and $V_0 \propto M^{\frac{1}{2}}$. When $M$ is small, $(1-\beta) \propto M^2$, and $V_0 \propto M^{-\frac{1}{2}}$. Thus $V_0$ decreases with increasing $M$ when $M$ is small, but increases when $M$ is large. Consequently $V_0$ must have a minimum for some value of $M$.

To find the minimum, put $\qquad M/\odot = x, \quad \dfrac{M/\odot}{1-\beta} = y,$

so that $\qquad 1-\beta = x/y.$

Equation (96) then gives $\qquad \dfrac{x}{y} = 0\cdot0026 x^2 m'^4 \left(1 - \dfrac{x}{y}\right)^4.$

Differentiating with regard to $x$ and putting $dy/dx = 0$, we find that the minimum of $y$ (*i.e.* of $V_0$ and $q_0$) occurs for

$$x/y = 1 - \beta = \tfrac{1}{5},$$

whence

$$\frac{M}{\odot} = \frac{1}{m'^2 \beta^2} \left(\frac{1-\beta}{0\cdot0026}\right)^{\frac{1}{2}} = \frac{13\cdot7}{m'^2} \qquad \text{(100)}.$$

This gives the mass for which, amongst giants of given temperature, the surface value of the gravitational potential is a minimum.

For Eddington's two values for the mean molecular weight throughout the star, $m' = 2\cdot8$ and $m' = 4$, we find $M = 1\cdot7\odot$, $M = 0\cdot86\odot$. These are not only of the order of magnitude of stellar masses, they are close to the actual masses of the majority of the stars. We thus have the following interesting result. *The surface value of the gravitational potential $GM/r_0$ considered as a function of the mass $M$ for a hypothetical series of stars of constant temperature tends to $+\infty$ for very small and very large masses, and has a single minimum; the existing stars are clustered about this minimum.* This result is perhaps little other than an alternative form of Eddington's result that in the neighbourhood of the masses of the existing stars $1-\beta$ changes rapidly from a value near zero (for smaller masses) to a value near unity (for larger masses). Whatever the mean molecular weight the minimum occurs when radiation pressure is $\frac{1}{5}$ of the total pressure. The mass for which the minimum occurs varies inversely as the square of the molecular weight, and the gravitational potential at the minimum as the inverse first power.

The tables given later show that the minimum is not a very sharp one.

The fact that most stars have the mass most favourable for loss by diffusion, combined with their permanency, suggests of itself that the loss by diffusion is always very small, as we shall soon see to be the case. It is perhaps tempting to suppose that the reason for the existence of the stars near this minimum is that they have been built up by a process of aggregation by capture—whilst the mass is still small every addition decreases the surface potential and so diminishes the power of making further captures. But the speculation seems neither profitable nor plausible.

§ 13. *Numerical application to stars.* Tables II—VI are based on Eddington's theory of the internal constitution of the stars. Table II gives the effective temperatures ($T_1$) and radii ($r_0$) of stars of given mass ($M$, here expressed in terms of the sun's mass) and mean density ($\rho_m$); it is an amplification of one given by Eddington*, and assumes a mean molecular weight $m' = 4$, with $k = 42{\cdot}5$ as the corresponding value of the absorption coefficient in the interior; the radii are given in cms. Table III gives the surface values of $g$ for the same stars, and also the values of $q_0$ *calculated for atomic hydrogen with* $T_0 = 2^{-\frac{1}{4}} T_1$. For a gas other than hydrogen the tabulated values must be multiplied by the molecular weight. Table IV gives the values of $\rho$ and $l$ at the escape-level for atomic hydrogen for the same stars. For this purpose the diameter of the hydrogen atom is taken to be that of the $2_2$ Bohr orbit (the second circular orbit), namely $4{\cdot}22 \times 10^{-8}$ cms., since the intensity of the Balmer spectrum at the temperatures concerned shows that an appreciable fraction of the atoms is in this state. However the orders of magnitude alone are material or indeed have any significance. The values of $\rho$ and $l$ are calculated from formulae (90) and (92) in which for definiteness $\mu_1$ has been taken to be $\frac{1}{4}$, $(\theta_1 = 75\frac{1}{2}^\circ)$, corresponding strictly to $n = \frac{2}{3}$.

The theoretical minimum for $q_0$ is not well indicated by these tables. Accordingly Tables V and VI have been calculated, for giant stars only. Table V shows the values of $r_0$, $g_0$, $q_0$, $\rho$, $l$ for giant stars of various masses at a constant effective temperature of 3000°, for a mean molecular weight $m' = 4$, $(k = 42{\cdot}5)$. Table VI is a similar table for $m' = 2{\cdot}8$, $(k = 23)$. The minimum of $q_0$ is apparent in each table. Values of the various quantities for other effective temperatures are deducible by using the relations

$$r_0 \propto T_1^{-2}, \quad g_0 \propto T_1^{4}, \quad q_0 \propto T_1, \quad \rho \propto T_1^{3}, \quad l \propto T_1^{-3}.$$

(These relations hold only for giant stars of constant mass.)

The minimum value of $q_0$ occurring in Table V is 277, $(M = 0{\cdot}855\odot)$. Stars of all other masses and of all higher effective temperatures (of mean molecular weight 4) yield theoretically higher values of $q_0$ than this; moreover since this value refers to atomic hydrogen, all other gases will again yield higher values. Thus out of all stars of effective temperature 3000° and higher, giant stars of mass $0{\cdot}855\odot$ and effective temperature 3000° will lose atomic hydrogen more

TABLE II. *Effective temperatures and radii* (*cms.*).

| $\rho_m$ | $M=0{\cdot}5$ | | $M=1{\cdot}0$ | | $M=1{\cdot}5$ | | $M=4{\cdot}5$ | |
|---|---|---|---|---|---|---|---|---|
| | $T_1$ | $r_0$ | $T_1$ | $r_0$ | $T_1$ | $r_0$ | $T_1$ | $r_0$ |
| 1·94 | 2680 | $0{\cdot}49 \times 10^{11}$ | 4060 | $0{\cdot}62 \times 10^{11}$ | 5100 | $0{\cdot}71 \times 10^{11}$ | 9590 | $1{\cdot}02 \times 10^{11}$ |
| 1·378 (☉) | | | 5860 | 0·696 | | | | |
| 1·12 | 4630 | 0·59 | 6930 | 0·745 | 8670 | 0·85 | 15050 | 1·23 |
| 0·64 | 5970 | 0·71 | 8770 | 0·90 | 10780 | 1·03 | 16780 | 1·48 |
| 0·356 | 6710 | 0·87 | 9710 | 1·09 | 11590 | 1·25 | 16680 | 1·80 |
| 0·198 | 7070 | 1·05 | 9930 | 1·33 | 11610 | 1·52 | 15880 | 2·19 |
| 0·106 | 7070 | 1·30 | 9680 | 1·64 | 11140 | 1·87 | 14760 | 2·70 |
| 0·055 | 6790 | 1·62 | 9120 | 2·03 | 10380 | 2·33 | 13480 | 3·36 |
| 0·0066 | 5410 | 3·28 | 6980 | 4·13 | 7800 | 4·72 | 9820 | 6·81 |
| 0·001 | 3950 | 6·14 | 5100 | 7·73 | 5700 | 8·85 | 7160 | 12·8 |
| 0·0002 | 3020 | 10·5 | 3900 | 13·2 | 4360 | 15·1 | 5480 | 21·8 |
| 0·0001 | 2690 | $13{\cdot}2 \times 10^{11}$ | 3480 | 16·7 | 3880 | 19·1 | 4870 | 27·5 |
| 0·00005 | | | 3100 | 21·0 | 3470 | 24·1 | 4370 | 34·7 |
| 0·00002 | | | 2660 | $28{\cdot}5 \times 10^{11}$ | 2970 | 32·6 | 3730 | 47·0 |
| 0·00001 | | | | | 2640 | $41{\cdot}2 \times 10^{11}$ | 3320 | $59{\cdot}4 \times 10^{11}$ |

* "Das Strahlungsgleichgewicht der Sterne," *Zeit. für Phys.*, Bd. 7, S. 381 (1921).

TABLE III. *Values of $g_0$ and $q_0$. (Atomic hydrogen.)*

| $\rho_m$ | $M=0\cdot5$ $g_0$ | $M=0\cdot5$ $q_0$ | $M=1\cdot0$ $g_0$ | $M=1\cdot0$ $q_0$ | $M=1\cdot5$ $g_0$ | $M=1\cdot5$ $q_0$ | $M=4\cdot5$ $g_0$ | $M=4\cdot5$ $q_0$ |
|---|---|---|---|---|---|---|---|---|
| 1·94 | $2\cdot66 \times 10^4$ | 7050 | $3\cdot35 \times 10^4$ | 7400 | $3\cdot84 \times 10^4$ | 7710 | $5\cdot57 \times 10^4$ | 8540 |
| 1·378(⊙) | | | 2·67 ,, | 4560 | | | | |
| 1·12 | 1·94 ,, | 3400 | 2·33 ,, | 3600 | 2·65 ,, | 3780 | 3·85 ,, | 4530 |
| 0·64 | 1·27 ,, | 2190 | 1·60 ,, | 2370 | 1·84 ,, | 2520 | 2·65 ,, | 3380 |
| 0·356 | $8\cdot60 \times 10^3$ | 1610 | 1·09 ,, | 1760 | 1·24 ,, | 1930 | 1·80 ,, | 2790 |
| 0·198 | 5·81 ,, | 1250 | $9\cdot72 \times 10^3$ | 1410 | $8\cdot39 \times 10^3$ | 1580 | 1·21 ,, | 2400 |
| 0·106 | 3·83 ,, | 1020 | 4·83 ,, | 1200 | 5·51 ,, | 1340 | $8\cdot00 \times 10^3$ | 2110 |
| 0·055 | 2·47 ,, | 850 | 3·12 ,, | 1010 | 3·56 ,, | 1153 | 5·16 ,, | 1850 |
| 0·0066 | $6\cdot00 \times 10^2$ | 526 | $7\cdot60 \times 10^2$ | 646 | $8\cdot70 \times 10^2$ | 756 | 1·25 ,, | 1250 |
| 0·001 | 1·71 ,, | 384 | 2·16 ,, | 471 | 2·47 ,, | 554 | $3\cdot57 \times 10^2$ | 915 |
| 0·0002 | $5\cdot85 \times 10$ | 294 | $7\cdot40 \times 10$ | 362 | $8\cdot47 \times 10$ | 423 | 1·22 ,, | 701 |
| 0·0001 | 3·7 ,, | 261 | 4·65 ,, | 321 | 5·30 ,, | 377 | $7\cdot70 \times 10$ | 625 |
| 0·00005 | | | 2·92 ,, | 285 | 3·15 ,, | 335 | 4·84 ,, | 553 |
| 0·00002 | | | 1·59 ,, | 246 | 1·82 ,, | 288 | 2·64 ,, | 478 |
| 0·00001 | | | | | 1·14 ,, | 256 | 1·65 ,, | 425 |

TABLE IV. *Values of $\rho$ and $l$ (cms.). (Atomic hydrogen.)*

| $\rho_m$ | $M=0\cdot5$ $\rho$ | $M=0\cdot5$ $l$ | $M=1\cdot0$ $\rho$ | $M=1\cdot0$ $l$ | $M=1\cdot5$ $\rho$ | $M=1\cdot5$ $l$ | $M=4\cdot5$ $\rho$ | $M=4\cdot5$ $l$ |
|---|---|---|---|---|---|---|---|---|
| 1·94 | $1\cdot1 \times 10^{-17}$ | $2\cdot0 \times 10^7$ | $8\cdot7 \times 10^{-18}$ | $2\cdot4 \times 10^7$ | $8\cdot0 \times 10^{-18}$ | $2\cdot6 \times 10^7$ | $6\cdot1 \times 10^{-18}$ | $3\cdot4 \times 10^7$ |
| 1·378(⊙) | | | 4·8 ,, | 4·3 ,, | | | | |
| 1·12 | $4\cdot2 \times 10^{-18}$ | 4·9 ,, | 3·5 ,, | 5·9 ,, | 3·3 ,, | 6·4 ,, | 2·7 ,, | 7·7 ,, |
| 0·64 | 2·2 ,, | 9·2 ,, | 1·9 ,, | $1\cdot1 \times 10^8$ | 1·8 ,, | $1\cdot1 \times 10^8$ | 1·7 ,, | $1\cdot2 \times 10^8$ |
| 0·356 | 1·4 ,, | $1\cdot5 \times 10^8$ | 1·2 ,, | 1·7 ,, | 1·1 ,, | 1·8 ,, | 1·1 ,, | 1·8 ,, |
| 0·198 | $8\cdot7 \times 10^{-19}$ | 2·4 ,, | $7\cdot8 \times 10^{-19}$ | 2·6 ,, | $7\cdot6 \times 10^{-19}$ | 2·7 ,, | $8\cdot0 \times 10^{-19}$ | 2·6 ,, |
| 0·106 | 5·7 ,, | 3·6 ,, | 5·4 ,, | 3·8 ,, | 5·2 ,, | 4·0 ,, | 5·7 ,, | 3·6 ,, |
| 0·055 | 3·8 ,, | 5·4 ,, | 3·6 ,, | 5·7 ,, | 3·6 ,, | 5·7 ,, | 4·1 ,, | 5·0 ,, |
| 0·0066 | 1·2 ,, | $1\cdot8 \times 10^9$ | 1·1 ,, | $1\cdot8 \times 10^9$ | 1·2 ,, | $1\cdot8 \times 10^9$ | 1·3 ,, | $1\cdot6 \times 10^9$ |
| 0·001 | $4\cdot6 \times 10^{-20}$ | 4·5 ,, | $4\cdot5 \times 10^{-20}$ | 4·6 ,, | $4\cdot6 \times 10^{-20}$ | 4·5 ,, | $5\cdot3 \times 10^{-20}$ | 3·9 ,, |
| 0·0002 | 2·1 ,, | $1\cdot0 \times 10^{10}$ | 2·0 ,, | $1\cdot0 \times 10^{10}$ | 2·0 ,, | $1\cdot0 \times 10^{10}$ | 2·4 ,, | 8·8 ,, |
| 0·0001 | 1·5 ,, | 1·4 ,, | 1·4 ,, | 1·5 ,, | 1·4 ,, | 1·4 ,, | 1·7 ,, | $1\cdot2 \times 10^{10}$ |
| 0·00005 | | | $9\cdot9 \times 10^{-21}$ | 2·1 ,, | 1·0 ,, | 2·0 ,, | 1·2 ,, | 1·8 ,, |
| 0·00002 | | | 6·3 ,, | 3·3 ,, | $6\cdot5 \times 10^{-21}$ | 3·2 ,, | $7\cdot5 \times 10^{-21}$ | 2·8 ,, |
| 0·00001 | | | | | 4·6 ,, | 4·5 ,, | 5·2 ,, | 4·0 ,, |

TABLE V. $T_1=3000°$, $m'=4$, $k=42.5$. (*Giant stars.*)

| $M$ | $r_0$ | $g_0$ | $q_0$ | $\rho$ | $l$ |
|---|---|---|---|---|---|
| 0·3 | $0.56\times10^{12}$ | 124 | 332 | $4.3\times10^{-20}$ | $4.8\times10^{9}$ |
| 0·5 | 1·06 ,, | 57·3 | 291 | 2·0 ,, | $1.0\times10^{10}$ |
| 0·855 | 1·91 ,, | 30·4 | 277 | 1·1 ,, | 2·0 ,, |
| 1·0 | 2·22 ,, | 26·2 | 278 | $9.2\times10^{-21}$ | 2·3 ,, |
| 1·5 | 3·19 ,, | 19·0 | 291 | 6·7 ,, | 3·1 ,, |
| 3·0 | 5·48 ,, | 12·9 | 339 | 4·5 ,, | 4·6 ,, |
| 4·5 | 7·33 ,, | 10·8 | 381 | 3·8 ,, | 5·5 ,, |
| 6·5 | 9·23 ,, | 9·9 | 436 | 3·5 ,, | 6·0 ,, |
| 9·0 | 11·30 ,, | 9·1 | 492 | 3·2 ,, | 6·5 ,, |
| $\infty$ | $\infty$ | 6·1 | $\infty$ | 2·1 ,, | 9·7 ,, |

TABLE VI. $T_1=3000°$, $m'=2.8$, $k=23$. (*Giant stars.*)

| $M$ | $r_0$ | $g_0$ | $q_0$ | $\rho$ | $l$ |
|---|---|---|---|---|---|
| 0·5 | $0.84\times10^{12}$ | 91·6 | 369 | $3.2\times10^{-20}$ | $6.5\times10^{9}$ |
| 1·0 | 2·03 ,, | 31·1 | 304 | 1·1 ,, | $1.9\times10^{10}$ |
| 1·5 | 3·19 ,, | 19·0 | 291 | $6.7\times10^{-21}$ | 3·1 ,, |
| 1·71 | 3·65 ,, | 16·5 | 290 | 5·8 ,, | 3·6 ,, |
| 3·0 | 6·11 ,, | 10·3 | 303 | 3·6 ,, | 5·7 ,, |
| 4·5 | 8·46 ,, | 8·1 | 328 | 2·8 ,, | 7·3 ,, |
| 6·5 | 11·12 ,, | 6·8 | 361 | 2·4 ,, | 8·7 ,, |
| 9·0 | 14·06 ,, | 5·9 | 396 | 2·1 ,, | $1.0\times10^{11}$ |
| $\infty$ | $\infty$ | 3·3 | $\infty$ | 1·2 ,, | 1·8 ,, |

rapidly than any other star will lose hydrogen or any other gas. The greatest rate of loss occurs for $n=0$. Inserting in formula (93) the rate of loss in grams per second is found to be

$$2\times10^{-87}\rho_0,$$

*i.e.* in grams per $10^6$ years $\qquad 6\times10^{-74}\rho_0.$

Whatever permissible value is adopted for the limb-density $\rho_0$ (see § 14), the rate of loss is completely negligible; more than this, scarcely any molecule ever escapes. The result is unaltered if we adopt the highest possible surface temperature, $T_0=T_1$; the index in the last expression merely becomes 56.

For giant stars at 3000° the value of $\rho$ for atomic hydrogen is of the order of magnitude of $10^{-20}$ gram. cm.$^{-3}$, corresponding to about $10^4$ atoms per cm.$^3$ and a mean free path of the order of $10^5$ kms. These are the critical density and mean free path at which the atmosphere ceases to behave as a gas in the ordinary sense and becomes a collection of molecules projected under gravity and continually falling back (since escape is negligible) into the denser atmosphere. For dwarf stars at the temperature of the sun, $\rho$ is of the order of $10^{-18}$ gram. cm., corresponding to about $10^6$ atoms per cm.$^3$ and to a mean free path of about 400 kms.

§ 14. *The height of the escape-level, and the density at the limb of a star.* The height above the limb at which escape (if any) may be said to be occurring is given by (91). It depends on $\rho_0$, the partial density at the limb of the constituent considered. Owing to the wide variations in estimates of the pressure in the atmosphere of a star and the difficulty of assigning precise levels to particular pressures, it is impossible to estimate $\rho_0$ with any accuracy; but the resulting uncertainty in $r/r_0$ is much smaller, since $\rho_0$ occurs as a logarithm.

Whatever estimate is made of $\rho_0$, it is clearly desirable to make consistent estimates for stars of different masses and temperatures. We will therefore attempt to discuss the variation of $\rho_0$. If $O$ denotes an observer, $C$ the centre of the star and $OCP$ a right angle, then $P$ will appear to be at the limb of the star provided the total optical thickness measured from $P$ along $PO$ just reaches, without exceeding, a certain value. Assume that the mean coefficient of absorption is roughly the same in all stellar atmospheres. Then if $t$ denotes the distance of any point on $OP$ from $P$, the position of $P$ when at the limb is determined by an equation of the form

$$\int_0^\infty \rho(r)\,dt = \text{const.} \qquad (101),$$

where

$$r^2 = r_0^2 + t^2.$$

The value of $\rho$ as a function of $t$ can be inserted in (101), and an asymptotic approximation to the integral obtained by the methods of Section II. Relation (101) is then found to reduce to

$$\rho_0 r_0 q_0^{-\frac{1}{2}} = \text{const.} \qquad (102),$$

whatever the value of $n$. Assuming that the mean molecular weight $m'$ is constant, this may also be written in either of the forms

$$\rho_0 \propto r_0^{-\frac{3}{2}} T_0^{-\frac{1}{2}} M^{\frac{1}{2}} \qquad (103),$$

$$\rho_0 \propto \rho_m^{\frac{1}{2}} T_0^{-\frac{1}{2}} \qquad (104),$$

$\rho_m$ being the mean density of the star*.

Equations (102) or (104) give the variation of the density at the limb for stars of different masses, radii and temperatures. For example, for a giant and dwarf of the same mass and temperature, we find

$$\frac{\rho_0\,(\text{giant})}{\rho_0\,(\text{dwarf})} = \left(\frac{\rho_m\,(\text{giant})}{\rho_m\,(\text{dwarf})}\right)^{\frac{1}{2}}.$$

Again, (104) shows that during the evolution of a giant $\rho_0 \propto \rho_m^{5/12}$, on Eddington's theory.

For definiteness we will take the partial pressure of atomic hydrogen at the limb of the sun to be $10^{-1}$ atmos.; taking the temperature there to be $5860° \times 2^{-\frac{1}{4}} = 4930°$, we find $\rho_0 = 2{\cdot}5 \times 10^{-7}$. The corresponding values of $\rho_0$ for other stars are shown in the second column of Table VII; they have been deduced by using formula (102). (Formula (102) strictly applies only to the total density at the limb, but in the absence of more definite knowledge we assume it to apply to partial densities.) For brevity Table VII has been calculated only for stars of the mass of the sun, but the results for other masses are of the same order of magnitude. The third column gives the fraction of the radius beyond the limb to which the atmosphere extends, calculated from (91); and the fourth column gives the actual heights in cms. If the limb-density on the sun had been

* Equation (102) should be compared with the corresponding equation usually given for the density $\rho_0'$ at the photosphere at the centre of the disc. Assuming this is determined by a relation of the form "optical thickness = const.," we find $k\rho_0/g$ = const., whence if $k$ is constant

$$\rho_0' r_0 q_0^{-1} = \text{const.}$$

This may also be written in the form

$$\rho_0' \propto \rho_m r_0 T_0^{-1}.$$

It may be mentioned here that calculations of this kind appear to show that the "mean density of the atmosphere above the visible surface" is not a function of the mean density of the star only (for stars of given temperature), but is also a function of the radius; it is more accurately regarded as a function of $g$. The contrary is however assumed by Russell in the explanation offered by him of the apparently small dispersion of mass amongst dwarf stars as deduced from spectroscopic parallaxes (*Astrophys. Journ.* 55, 239, 1922).

TABLE VII. *Limb-densities and surface heights* (*cms.*). ($M = 1$; *atomic hydrogen.*)

| $\rho_m$ | $\rho_0$ | $\frac{r - r_0}{r_0}$ | $r - r_0$ |
|---|---|---|---|
| 1·94 | $3·5 \times 10^{-7}$ | 0·00033 | $2·0 \times 10^{8}$ |
| 1·378 (☉) | 2·5 ,, | 0·00054 | 3·8 ,, |
| 1·12 | 2·0 ,, | 0·00069 | 5·1 ,, |
| 0·64 | 1·3 ,, | 0·00105 | 9·5 ,, |
| 0·356 | $9·7 \times 10^{-8}$ | 0·0014 | $1·6 \times 10^{9}$ |
| 0·198 | 7·1 ,, | 0·0018 | 2·4 ,, |
| 0·106 | 5·4 ,, | 0·0021 | 3·5 ,, |
| 0·055 | 4·0 ,, | 0·0025 | 5·1 ,, |
| 0·0066 | 1·5 ,, | 0·0040 | $1·6 \times 10^{10}$ |
| 0·001 | $7·1 \times 10^{-9}$ | 0·0055 | 4·2 ,, |
| 0·0002 | 3·7 ,, | 0·0072 | 9·5 ,, |
| 0·0001 | 2·7 ,, | 0·0081 | $1·4 \times 10^{11}$ |
| 0·00005 | 2·0 ,, | 0·0092 | 1·9 ,, |
| 0·00002 | 1·4 ,, | 0·0106 | 3·0 ,, |
| 0·00001 | 1·0 ,, | 0·0120 | 4·3 ,, |

taken to correspond to a partial pressure of $10^{-4}$ atmos., the numbers in the third and fourth columns would require to be multiplied by a factor of about $\frac{2}{3}$.

It appears that on the sun the escape-level for atomic hydrogen should be reached at 3000—4000 kms. above the limb; for a giant $M$ star of the same mass, at about $10^5$ kms. For heavier elements these numbers would have to be reduced considerably, according to the value of $q_0$. It is not suggested that the heights thus found should correspond in any way with observation; they are merely of interest as being the heights implied by a purely gravitational theory of equilibrium. The inadequacy of such a theory to account for the existence of the sun's chromosphere is well known.

§ 15. *The escape of electrons.* It has been seen (at least on simple gravitational equilibrium) that the loss due to the escape of hydrogen atoms is negligible, and a fortiori that due to all heavier atoms. But in the atmosphere of a star, owing to ionization, there will be an abundance of free electrons. The mass of an electron being 1/1835 of that of a hydrogen atom, the value of $q_0$ for electrons is 1/1835 that for atomic hydrogen. Reference to Table III shows that for electrons $q_0$ will vary from about 0·1 for giants to about 4 for dwarfs. For values of $q_0$ so small as this the rate of escape will be very appreciable.

But the escape of electrons, however large initially, cannot continue indefinitely, for it will be checked by the attraction exerted on the electrons by the steadily increasing positive charge acquired by the star. Escape will thus continually slacken, until a state is reached in which further loss is inappreciable. The star will then have a permanent positive charge.

To estimate the order of magnitude of this charge, suppose that at any particular stage in the evaporation $N$ electrons have escaped, of total charge $Ne = E$. The star will then have a positive charge of the same amount, and this charge will make itself apparent as an excess of positively charged ions. Owing to the high degree of ionization, the star will behave as a conductor, and the charge will distribute itself over the surface. Thus the excess of ionized atoms will be found chiefly in what we have called the surface region of the atmosphere. The equilibrium of the lower parts of the atmosphere will therefore not be affected, but the presence of the charge will make the outer parts expand slightly. We shall not here discuss the nature of the new state of equilibrium, but it is clear that immediately outside the surface region the external force acting on an electron of mass $m_e$ will be

$$\frac{Gm_e M}{r^2} - \frac{eE}{r^2},$$

and that on a singly ionized atom of mass $m_i$

$$\frac{Gm_i M}{r^2} - \frac{eE}{r^2}.$$

Hence the value of $q$ for an electron is

$$\frac{Gm_e M}{rRT} - \frac{eE}{rRT},$$

and for an ion

$$\frac{Gm_i M}{rRT} - \frac{eE}{rRT}.$$

Thus in each case $eE/rRT$ gives the correction to $q$ due to the charge on the star. When the star has become charged, the gravitational part $Gm_e M/rRT$ for an electron is negligible compared with the electrostatic part.

Since the state of equilibrium below the surface region is different from that in the surface region, owing to the electrostatic forces, our detailed formulae for the rate of loss will not apply directly to the electronic loss; also two kinds of atoms are present—ions and electrons. However we can apply formula (30). From this the charge gained per second due to the escape of electrons is practically

$$\frac{dE}{dt} = 2n_1 e r_0^2 \left(\frac{2\pi RT_0}{m_e}\right)^{\frac{1}{2}} e^{-q}(1+q),$$

where $q$ has the above value and $n_1$ is the number of electrons per unit volume on the outer fringe of the surface region. This may be written approximately

$$\frac{1}{E}\frac{dE}{dt} = \frac{2n_1 e^2 r_0}{RT_0}\left(\frac{2\pi RT_0}{m_e}\right)^{\frac{1}{2}} e^{-eE/r_0 RT_0}.$$

It is not easy to estimate $n_1$. But we saw in § 13 that the number of hydrogen atoms per unit volume in the surface region of a giant star is about $10^4$, and taking into account the degree of ionization the number of free electrons from all sources must be at least $10^3$ per unit volume. Applying the formula now to a giant $M$ star of the mass of the sun, for which $T = 3000°$, $T_0 = 2520°$, $r_0 = 2{\cdot}2 \times 10^{12}$, we find

$$\frac{1}{E}\frac{dE}{dt} = e^{-eE/r_0 RT_0} \times 10^{17}.$$

Considered as a differential equation, this formula implies that $E$ increases indefinitely with $t$. But just as the escape of gas molecules is completely negligible, so the further escape of electrons will be completely negligible, even in periods of time astronomically large, once $E$ has attained a certain value. When the charge acquired is $E$, the further charge $\delta E$ acquired in the next $10^5$ years is given roughly by

$$\frac{\delta E}{E} = 10^{29} \times e^{-eE/r_0 RT_0} = 10^{29} \times 10^{-0{\cdot}434 eE/r_0 RT_0}.$$

In order that this may be negligible, $0{\cdot}434 eE/r_0 RT_0$ must be at least about 30. As regards order of magnitude, then, we see that the escape of electrons will endow the star with a positive charge $E$ which will be at least as large as is given by

$$0{\cdot}434 eE/r_0 RT_0 = 30,$$

but which cannot seriously exceed this value. From this equation we find

$$eE/r_0 RT_0 = 70, \quad E/r_0 = 5 \times 10^{-2} \text{ E.S.U.} = 15 \text{ volts}, \quad E = 1{\cdot}1 \times 10^{11} \text{ E.S.U.}$$

It is clear that these estimates are very little affected by the uncertainty in the value of $n_1$.

Thus for a giant star of type $M$ the value of $q_0$ for ionized hydrogen must be taken to be smaller by about 70 than the value calculated on the simple gravitational theory. The latter has been seen to be about 278. Hence the effect of the charge acquired by the escape of electrons is a force of repulsion, on ionized hydrogen atoms, equal to about $\frac{1}{4}$ of the force of attraction due to gravity. Theoretically this facilitates the escape of positively charged hydrogen atoms. If this reduction in the effective value of gravity were to permit an appreciable rate of escape of ionized hydrogen, the state would not be permanent, as the positive charge would tend to be dissipated. But it is easily seen from the calculations of § 13 that the rate of escape of hydrogen is still completely negligible; and on heavier atoms the effect will be still less, for the electrostatic force will be the same whilst the force due to gravity varies as the mass.

If we make similar calculations for a dwarf star such as the sun, the value of $eE/r_0 RT_0$ comes out about the same. Putting $T_0 = 5860^\circ \times 2^{-\frac{1}{4}} = 4930^\circ$, $r_0 = 7 \times 10^{10}$, we find

$$E/r_0 = 30 \text{ volts}, \quad E = 0{\cdot}7 \times 10^{10} \text{ E.S.U.}$$

Owing to the smaller radius the charge is about $\frac{1}{15}$ that for the giant.

But if the sun has evolved from a giant star of type $M$, it should in that state have possessed a charge of $1{\cdot}1 \times 10^{11}$ E.S.U. It could only in its present state have a smaller charge (which has been calculated as though the sun had always been in its present state) if it has lost positive charge during its evolution from the $M$ stage onwards. But we have seen that the escape of positively charged hydrogen ions is negligible, and so far as the range of phenomena here considered is concerned this is the only way in which a positive charge can be lost—unless indeed a star can in a later stage of evolution pick up again out of space the electrons it lost in an earlier stage. Assuming the latter is impossible it follows that the sun must still have the charge of $1{\cdot}1 \times 10^{11}$ E.S.U. which it had as a giant $M$ star. This makes its present potential equal to 470 volts, and its present value of $eE/r_0 RT_0$ equal to 1100. The latter number, therefore, is the amount by which the calculated value of $q$ must be diminished, in the case of charged atoms, to allow for electrostatic repulsion. For hydrogen ions we had $q = 4560$, so that as for a giant about $\frac{1}{4}$ of the weight of hydrogen ions is supported by electrostatic repulsion. It is clear, in fact, that charge and mass both remaining constant during evolution, the ratio of the electrical and gravitational forces will remain constant. The potential, however, will steadily increase during evolution, since the electrostatic capacity decreases.

Various phenomena of solar and geo-physics have suggested theories which demand the emission from time to time of charged particles from the sun—some theories demanding positive charges only, some negative. It is possible too that electrons or ions or both are discharged from the sun at the vertices of prominences, either eruptively or owing to the discharging action of a point. If such discharges occur, the above speculation concerning the result of supposing the charge acquired as an $M$ star to be conserved becomes of no importance. But we can assign limits between which the potential of the sun must lie whenever it is in a quiescent non-emitting state. If by the emission of positively charged particles the sun acquires a negative potential, or a positive potential appreciably less than the value 30 volts calculated above, then due merely to thermal agitation electrons will escape from the surface until a positive potential of about 30 volts is restored. If by the emission of electrons due to some cause other than thermal agitation the sun acquires a positive potential in excess of that capable of just retaining hydrogen ions, then hydrogen ions will escape until the positive potential is reduced to the latter value.

To calculate this maximum positive potential in the steady state, if $E$ is the positive charge at any moment the rate of loss is

$$-\frac{dE}{dt} = 2n_1 e r_1^2 \left(\frac{2\pi R T_0}{m_i}\right)^{\frac{1}{2}} e^{-\frac{Gm_i M}{r_0 RT_0} + \frac{eE}{r_0 RT_0}} \left(\frac{Gm_i M}{r_0 RT_0} - \frac{eE}{r_0 RT_0}\right).$$

We know from Table III that $Gm_i M/r_0 RT_0 = 4560$. Determining $eE/r_0 RT_0$ so as to make the rate of loss inappreciable (it is almost immaterial whether we consider the loss in one year or in $10^5$ years) we find

$$0{\cdot}434\,(4560 - eE/r_0 RT_0) = 25,$$

$$eE/r_0 RT_0 = 4500,$$

$$E/r_0 = 6{\cdot}3 \text{ E.S.U.} = 1900 \text{ volts}.$$

(This estimate agrees with that of Lindemann*, who calculated it directly from the crude equilibrium relation $eE = Gm_i M$.) Thus when in a quiescent state the potential of the sun must lie between about 30 and 1900 volts. Notice that the "evolutionary" potential 470 volts lies between these limits, as it must.

The maximum positive potential that a giant star of type $M$ of the mass of the sun can possess without losing hydrogen ions may be calculated in the same way to be 44 volts; the minimum is the value 15 volts already calculated.

§ 16. *Application to the earth's atmosphere.* It has been shown by many writers that the earth's atmosphere at its present temperature retains hydrogen if hydrogen is an existing constituent, and a fortiori helium. But it is of interest to apply our more detailed formulae to the case of the earth.

We adopt $T_0 = 219°$, the mean temperature of the stratosphere (or at least the base of the stratosphere) over S.E. England; the value of $T_0$ varies considerably with latitude, but there are theoretical grounds based on radiation theory† for supposing that the mean temperature of the stratosphere over the whole earth is close to the value adopted. The suffix 0 is to refer to the base of the stratosphere. Taking $r_0 = 6{\cdot}38 \times 10^8$ cms., the value of $q_0$ for *molecular* hydrogen is 69·2 and for helium is 137. The density of helium at the base of the stratosphere (12 kms. height, say) is calculated‡ to be $1{\cdot}71 \times 10^{-10}$. If the determinations of hydrogen content in the lower atmosphere may be taken as reliably indicating an upper hydrogen atmosphere, then the density of this at the base of the stratosphere may be estimated at about $2{\cdot}1 \times 10^{-10}$ (corresponding to about 1 part in $10^5$ by volume). We then find the following (Table VIII), the calculation being made as before for $\mu_1 = \frac{1}{4}$, $\theta_1 = 75\frac{1}{2}°$, ($n = \frac{2}{3}$).

The helium results must be taken as giving the boundary of the atmosphere if hydrogen is absent. Assuming a helium atmosphere, we see that the surface region occurs at 630 kms.; above this height therefore, the molecules are chiefly in free flight without collisions. The smallness of this height and the smallness of the "free path" there (130 kms.) are perhaps surprising, considering the much larger estimates which have sometimes been made. It would be interesting

* *Phil. Mag.*, 38, 674, 1919.

† *Ibid.*, 44, 892, 1922.

‡ Taken from Chapman and Milne, *Quart. Journ. Roy. Met. Soc.*, 46, 370, 1920.

TABLE VIII.

| | Hydrogen | Helium |
|---|---|---|
| $q_0$ | 69·2 | 137 |
| $\rho_0$ | $2{\cdot}1\times10^{-10}$ gram cm.$^{-3}$ | $1{\cdot}71\times10^{-10}$ gram cm.$^{-3}$ |
| $\rho$ | $4{\cdot}9\times10^{-17}$ ,, ,, | $2{\cdot}3\times10^{-16}$ ,, ,, |
| $l/r_0$ | 0·041 | 0·021 |
| $l$ | 260 kms. | 130 kms. |
| $r/r_0$ | 1·22 | 1·099 |
| $r-r_0$ | 1400 kms. | 630 kms. |

to study this height in connection with the observed maximum heights of aurorae. It must be remembered that if the theory presented in this paper is correct, it should apply rigorously to the earth's atmosphere, for here there is no complication due to radiation pressure as in stellar atmospheres.

*Summary.*

§ 17. The paper first discusses the hydrostatics of a gaseous gravitating atmosphere in which the temperature falls off as the inverse $n$th power of the distance from the centre of the nucleus. Asymptotic formulae are found for the pressure and density at large distances from the nucleus, for the different cases that arise according to the value of $n$. (Section I.)

Section II discusses the phenomenon of the escape of molecules from the fringe of an atmosphere of the above type. The analysis is carried out with the help of the concept of the "free solid angle" at any given level. As viewed from a point at a sufficiently high level in the atmosphere, the sky must appear to be partially clear of molecules down to a certain zenith distance, and inside this region it is the solid angle actually unoccupied which is available for escaping molecules. A formal expression is obtained for the rate of escape across any level, in the form of the product of the free solid angle and a function of the density at the level. For one particular level this has a maximum, which in general may be taken to give the actual rate of escape, and this level may be taken in general as the layer from which escape is occurring. At the escape-level the sky is partially clear of molecules down to a zenith distance of about 75°, though the actual value depends somewhat on the value of $n$. Simple formulae are obtained for the density and equivalent mean free path at the escape-level, and for the height of the level. The latter depends on the density at the reference level in the lower atmosphere, but the former depend only on the temperature, the value of gravity at the surface and the diameter of the molecules. The actual rate of escape is proportional to $\rho_0^{1-n}$, where $\rho_0$ is the density at the reference level, and to $\sigma^{-2n}$, where $\sigma$ is the diameter of a molecule. The asymptotic formulae of Section I are required to secure the convergence of integrals which arise in the analysis of Section II. A summary of the mathematical results of Section II is given in § 11.

In Section III the results are applied to stellar atmospheres (assuming that the equilibrium of such atmospheres is determined solely by gravitation). It is shown that on Eddington's theory of the internal equilibrium of a star, out of all possible masses the masses of the existing stars are grouped about that which has the least gravitational potential at the surface in the giant stage, *i.e.* about the mass most favourable for loss by diffusion to be appreciable. In spite of this the loss is found to be completely negligible for all stars. Tables are given for the density, mean free path, etc., at the escape-level, for stars of various masses and temperatures. The density at the limb of a star, and its variation with stage of evolution, etc., are discussed.

The escape of electrons from the surface of a star is considered. It is shown in agreement with other writers that the potential of the sun must be positive when in a steady state; and that it cannot permanently exceed 1900 volts or be less than 30 volts, as escape of either hydrogen nuclei or electrons would soon reduce it to the first or second of these values respectively. If electrons could be lost only by diffusion, the potential would have the lower value, as calculated from the sun's present state; but if the sun has evolved from a giant *M* star, and has retained the charge it should then have acquired through the then freer escape of electrons, its present potential should be about 470 volts. The potential of a giant *M* star should be between 15 and 44 volts.

Some application is made to the earth's atmosphere. The escape-level for helium is about 630 kms., that for hydrogen 1400 kms., though the actual escape is negligible, as is well known. The corresponding mean free paths are about 130 kms., and 260 kms.

I wish to express my thanks to Mr W. H. Manning, of the Solar Physics Observatory, for assistance in the preparation of the diagrams.

# XXVII. *Some problems of Diophantine approximation: The analytic properties of certain Dirichlet's series associated with the distribution of numbers to modulus unity.*

By G. H. Hardy, Trinity College, and J. E. Littlewood, Trinity College.

[*Received* 6 June 1922.]

1·1. The series in question are

$$F_1(s)=\Sigma\frac{\alpha_n}{n^s},\quad F_2(s)=\Sigma\frac{\alpha_n^2-\frac{1}{12}}{n^s},\quad F_3(s)=\Sigma\frac{\alpha_n^3-\frac{1}{4}\alpha_n}{n^s},\quad \dots \qquad (1·11)$$

where
$$s=\sigma+it,\quad \alpha_n=\alpha_n(\theta)=\{n\theta\}=n\theta-[n\theta]-\tfrac{1}{2}, \qquad (1·111)$$
$\theta$ is irrational, $[x]$ is the integral part of $x$, and the summation (as always unless the contrary is stated) extends over positive integral values of $n$. The general formula for the $k$th function is

$$F_k(s)=F_k(s,\theta)=\Sigma\frac{\phi_k(n\theta)}{n^s}, \qquad (1·12)$$

where $\phi_k(x)$ is defined by

$$\phi_{2m}(x)=P_{2m}(x)+(-1)^{m-1}B_m,\quad \phi_{2m+1}(x)=P_{2m+1}(x),\quad (0<x<1) \qquad (1·131)$$

$$\phi(x+1)=\phi(x) \qquad (1·132)$$

and the $P$'s are Bernoulli's polynomials*.

The properties of these functions, which are very remarkable, are intimately bound up with the problem of the distribution of the numbers $n\theta$ to modulus 1†.

1·2. The properties of the function $F_1(s)$ have already been investigated by Hecke‡ when $\theta$ is a quadratic surd. Hecke supposes in particular that $\theta=\sqrt{D}$, where $D$ is free from squared factors and congruent to 2 or 3 to modulus 4. He shows that in this case $F_1(s)$ is meromorphic, and that its only possible singularities are simple poles at the points

$$s=-2q+\frac{2r\pi i}{\log\eta}, \qquad (1·21)$$

where
$$q=0,1,2,\dots;\ r=\dots,-1,0,1,2,\dots,\dots \qquad (1·211)$$
and $\eta$ is a particular unity of the corpus $K(\sqrt{D})$. His method rests upon the theory of the new 'Zeta-functions' which he has recently introduced into analysis, and there can be no doubt that it is the best for the particular problem with which he is concerned.

It is none the less of interest to discuss the function for general values of $\theta$, and by methods as elementary as possible. When we do this, we find ourselves compelled to treat $F_1(s)$ as the

* We follow the notation of Lindelöf (*Le calcul des résidus et ses applications à la théorie des fonctions*, 32 *et seq.*). The definition of the functions for integral values of $x$ is immaterial.

† In regard to this problem see the following memoirs:

G. H. Hardy and J. E. Littlewood, 'Some problems of Diophantine approximation': (1) *Proceedings of the Fifth International Congress of Mathematicians*, Cambridge, 1912, 1, 223—229; (2) 'The fractional part of $n^k\theta$', *Acta Math.*, 37 (1914), 155—190; (3) 'The lattice points of a right-angled triangle', *Proc. London Math. Soc.* (2), 20 (1921), 15—36; (4) 'The lattice points of a right-angled triangle (second memoir)', *Hamburg Math. Abh.*, 1 (1922), 212—249.

H. Weyl, 'Über die Gleichverteilung von Zahlen mod. Eins', *Math. Ann.*, 77 (1916), 313—352.

E. Hecke, 'Über analytische Funktionen und die Verteilung von Zahlen mod. Eins', *Hamburg Math. Abh.*, 1 (1921), 54—76.

A. Ostrowski; (1) 'Bemerkungen zur Theorie der Diophantischen Approximationen', *ibid.*, 77—98; (2) 'Zu meiner Note: Bemerkungen u.s.w.', *ibid.*, 250—251.

H. Behnke, 'Über die Verteilung von Irrationalitäten mod. 1', *ibid.*, 252—267.

‡ *l.c. supra.*

first of the sequence of functions $F_k(s)$. We also find ourselves led to the following classification of irrationals $\theta$.

We suppose, as we may without loss of generality, that $0<\theta<1$, and we write

$$\theta=\frac{1}{a_1+\theta_1},\quad \theta_1=\frac{1}{a_2+\theta_2},\quad \dots, \qquad (1\cdot22)$$

where $a_1, a_2, \dots$ are the partial quotients in the expression of $\theta$ as a simple continued fraction. We say that $\theta$ is *of class* $\lambda$ if $\lambda$ is the least number such that

$$(\theta\theta_1\dots\theta_{n-1})^{\lambda+\epsilon}/\theta_n\to 0 \qquad (1\cdot23)$$

for every positive $\epsilon$, or, what is the same thing, such that

$$n^{1+\lambda+\epsilon}\,|\sin n\theta\pi|\to\infty$$

for every positive $\epsilon$. If no such number exists, we say that $\theta$ is *of infinite class*. A quadratic surd is of class 0, and every algebraic number is of finite class.

Our principal results may be summarised as follows. In the first place, $F_k(s)$ *is regular for*

$$\sigma>\sigma_k=1-\frac{k}{1+\lambda}; \qquad (1\cdot24)$$

in particular, $F_1(s)$ is regular for $\sigma>\lambda/(1+\lambda)$. This we prove for $k>1$ in § 2, and for $k=1$ in § 3.

There are alternative proofs of this theorem. When $k=1$, it may be derived from Theorem 2 of our memoir (4), or from the sharper Theorem 5, due originally to Ostrowski; but the analysis of § 3 is necessary in any case for our further investigations. When $k>1$, it has been proved by Behnke*, by means of the formulae of linear transformation of the Theta-functions. The proof given here is a good deal simpler.

If $\lambda>0$, the result just stated is final; for then $\sigma=\sigma_k$ *is a singular line for the function.* We prove this in § 3. We have no doubt that the line is still singular when $\lambda=0$, *except when $\theta$ is quadratic*, so that the case considered by Hecke is completely exceptional; but this we are unable to prove.

In § 4 we consider the question of the convergence or summability of the series (1·11), and show that the regions of convergence or summability are always as extensive as is consistent with the analytic properties of the functions and the order of magnitude of the coefficients. Some theorems concerning convergence have been found already by Behnke†. These are included in ours, which assert the most that can be true.

2·1. THEOREM 1. *If $k>1$, and $\theta$ is of class $\lambda$, then $F_k(s)$ is regular for*

$$\sigma>\sigma_k=1-\frac{k}{1+\lambda}.$$

We have‡

$$\phi_{2m}(x)=\frac{(-1)^{m+1}2(2m)!}{(2\pi)^{2m}}\sum\frac{\cos 2\nu\pi x}{\nu^{2m}}\quad (m\geqq 1), \qquad (2\cdot111)$$

$$\phi_{2m+1}(x)=\frac{(-1)^{m+1}2(2m+1)!}{(2\pi)^{2m+1}}\sum\frac{\sin 2\nu\pi x}{\nu^{2m+1}}\quad (m\geqq 0). \qquad (2\cdot112)$$

It is therefore sufficient to show that the functions

$$g_k(s)=\sum\frac{\psi_k(n\theta)}{n^s},\quad h_k(s)=\sum\frac{\psi_k(-n\theta)}{n^s}, \qquad (2\cdot12)$$

where

$$\psi_k(x)=\sum\frac{e(\nu x)}{\nu^k}\quad (k>1)\S, \qquad (2\cdot121)$$

are regular for $\sigma>\sigma_k$. We shall discuss only $g_k(s)$, observing that our argument remains valid with a formal change throughout of $n\theta$ into $-n\theta$.

* Behnke, *l.c.*, 265—266. † Behnke, *l.c.*, 266. ‡ Lindelöf, *l.c.*, 34. § We write $e(x)$ for $e^{2\pi i x}$, following Weyl.

Suppose first that $\sigma > 1$. Then

$$g_k(s) = \sum_1^\infty \frac{1}{n^s} \sum_1^\infty \frac{e(\nu n\theta)}{\nu^k} = \sum_1^\infty \frac{\chi(\nu)}{\nu^k}, \quad \text{.............................(2·13)}$$

where

$$\chi(\nu) = \chi(s, \theta, \nu) = \sum_1^\infty \frac{e(\nu n\theta)}{n^s} \quad \text{.............................(2·131)}$$

This function is an integral function of $s$, and its continuation all over the plane is given by

$$\chi(\nu) = -\frac{\Gamma(1-s)}{2\pi i} \int_C \frac{e^{-x+2\nu\theta\pi i}}{1-e^{-x+2\nu\theta\pi i}} (-x)^{s-1}\, dx, \quad \text{.....................(2·14)}$$

where $C$ is a loop enclosing the positive real axis in the clockwise direction, and passing inside all the poles

$$x = x_m = 2\pi i(m + \nu\theta) \qquad (m = \ldots -1, 0, 1, \ldots)$$

of the integrand. We write

$$X(\nu) = X(s, \theta, \nu) = \frac{\chi(s, \theta, \nu)}{\Gamma(1-s)}, \quad G_k(s) = \frac{g_k(s)}{\Gamma(1-s)}. \quad \text{..................(2·15)}$$

2·2. There is one and only one of the numbers $x_m$ whose modulus is less than $\pi$. We define a number $\delta = \delta(\nu)$ as follows. If $\mathbf{x}_m$ is the $x_m$ of least modulus, and $|\mathbf{x}_m| \geqq \frac{1}{2}\pi$, we take $\delta = \frac{1}{4}\pi$. If $|\mathbf{x}_m| < \frac{1}{2}\pi$, we take $\delta = \frac{3}{4}\pi$. We denote by $C_0$ the contour formed by the semicircle $|x| = \delta$, $|\arg(-x)| \leqq \frac{1}{2}\pi$ and the two lines $\mathbf{R}(x) \geqq 0$, $|\mathbf{I}(x)| = \delta$. The distance of any point of $C_0$ from the nearest pole is greater than $\frac{1}{4}\pi$. Hence, if we write

$$X_0(\nu) = X_0(s, \theta, \nu) = \frac{1}{2\pi i} \int_{C_0} \frac{e^{-x+2\nu\theta\pi i}}{1-e^{-x+2\nu\theta\pi i}} (-x)^{s-1}\, dx, \quad \text{.................(2·21)}$$

we have

$$X_0(\nu) = O\left(\int_{C_0} |e^{-x}| \, |(-x)^{s-1}| \, |dx|\right) = O(1), \quad \text{.....................(2·22)}$$

uniformly throughout any bounded domain $T$ in the plane of $s$.

Now

$$X(\nu) = X_0(\nu) \quad (|\mathbf{x}_m| \geqq \tfrac{1}{2}\pi), \quad \text{.......................(2·231)}$$

and

$$X(\nu) = X_0(\nu) + (-\mathbf{x}_m)^{s-1} \quad (|\mathbf{x}_m| < \tfrac{1}{2}\pi). \quad \text{..............(2·232)}$$

The series

$$\sum \frac{X_0(\nu)}{\nu^k}$$

is, by (2·22), uniformly convergent throughout $T$, and its sum is an analytic function regular throughout $T$. It follows, from (2·13), (2·15), (2·231), and (2·232), that $G_k(s)$ is regular in any bounded domain throughout which the series

$$S = \sum \frac{(-\mathbf{x}_m)^{s-1}}{\nu^k} \quad (|\mathbf{x}_m| < \tfrac{1}{2}\pi) \quad \text{.....................(2·24)}$$

is uniformly convergent. This is certainly so if the series

$$\sum_1^\infty \frac{1}{\nu^k \, |\overline{\nu\theta}|^{1-\sigma}},$$

where $\overline{\nu\theta}$ is the difference between $\nu\theta$ and the integer nearest to $\nu\theta$, is uniformly convergent; and this series is, by Lemma 3 of our paper (4), uniformly convergent in any half-plane

$$\sigma \geqq 1 - \frac{k}{1+\lambda} + \epsilon > 1 - \frac{k}{1+\lambda}.$$

In other words, $G_k(s)$ is regular for the values of $s$ specified in the theorem.

It follows from (2·15) that $g_k(s)$ is also regular, except perhaps at the poles $s = 1, 2, 3, \ldots$ of $\Gamma(1-s)$. Of these, $s = 2, 3, \ldots$ are plainly not poles of $g_k(s)$. When $s = 1$, $X(\nu)$, and therefore $G_k(s)$, vanishes. Thus $g_k(s)$ is regular also for $s = 1$, which completes the proof of the theorem.

3·1. The method of § 2 fails when $k=1$, and more intricate analysis is necessary.

*Lemma A. If $\theta$ is irrational and positive, $x \geqq 0$, $y=\theta x$, and $f(0)=g(0)=0$, then*

$$\sum_{m \leqq x} f([m\theta])(g(m)-g(m-1)) + \sum_{n \leqq y} g\left(\left[\frac{n}{\theta}\right]\right)(f(n)-f(n-1)) = f([y])\, g([x]). \quad \text{.........(3·11)}$$

This is Lemma 7 of our paper (4).

*Lemma B. If $c>0$, $\xi$ is real, and*

$$c_1 = c/\theta, \quad \xi_1 = c-\theta\xi, \quad \beta_n = \{n/\theta\}, \quad \text{.........(3·12)}$$

*then*

$$\frac{e^{-\frac{1}{2}\xi}}{1-e^{-\xi}} \sum_1^\infty e^{-mc}(e^{\alpha_m \xi}-1) + \frac{e^{-\frac{1}{2}\xi_1}}{1-e^{-\xi_1}} \sum_1^\infty e^{-nc_1}(e^{\beta_n \xi_1}-1) = W, \quad \text{.........(3·13)}$$

*where*

$$W = W(c,\theta,\xi) = \frac{e^{-\xi-\xi_1}}{(1-e^{-\xi})(1-e^{-\xi_1})} - \frac{e^{-\frac{1}{2}\xi}}{1-e^{-\xi}}\,\frac{e^{-c}}{1-e^{-c}} - \frac{e^{-\frac{1}{2}\xi_1}}{1-e^{-\xi_1}}\,\frac{e^{-c_1}}{1-e^{-c_1}} \quad \text{...(3·131)}$$

In (3·11) take

$$f(u) = 1-e^{-u\xi}, \quad g(u) = 1-e^{-u\xi_1},$$

where $c=\theta\xi+\xi_1>0$, and make $x \to \infty$. We obtain

$$\frac{e^{-\xi}}{1-e^{-\xi}} \sum_1^\infty e^{-[m\theta]\xi - m\xi_1} + \frac{e^{-\xi_1}}{1-e^{-\xi_1}} \sum_1^\infty e^{-[n/\theta]\xi_1 - n\xi} = \frac{e^{-\xi-\xi_1}}{(1-e^{-\xi})(1-e^{-\xi_1})}.$$

Substituting for $[m\theta]$ and $[n/\theta]$ in terms of $\alpha_m$ and $\beta_n$, and making some simple reductions, we obtain (3·13).

Taking the limit of (3·13) as $\xi \to 0$, we obtain

*Lemma C. If $c>0$ and $c_1=c/\theta$, then*

$$\sum_1^\infty \alpha_m e^{-mc} + \frac{e^{-\frac{1}{2}c}}{1-e^{-c}} \sum_1^\infty (e^{\beta_n c}-1)e^{-nc_1} = w, \quad \text{.........(3·14)}$$

*where*

$$w = w(c,\theta) = \frac{\theta e^{-c}}{(1-e^{-c})^2} - \frac{1}{2}\,\frac{e^{-c}}{1-e^{-c}} - \frac{e^{-\frac{1}{2}c}}{1-e^{-c}}\,\frac{e^{-c_1}}{1-e^{-c_1}}. \quad \text{.........(3·141)}$$

*Lemma D. We have*

$$\sum_1^\infty \alpha_m e^{-mc} + \sum_1^\infty \beta_n e^{-nc_1} + \tfrac{1}{2}c \sum_1^\infty (\beta_n^2 - \tfrac{1}{12}) e^{-nc_1} = v + O(ce^{-c}) \quad \text{.........(3·15)}$$

*for all positive values of $c$, where*

$$v = v(c,\theta) = \frac{\theta e^{-c}}{(1-e^{-c})^2} - \frac{1}{2}\,\frac{e^{-c}}{1-e^{-c}} - \frac{1}{c}\,\frac{e^{-c_1}}{1-e^{-c_1}}. \quad \text{.........(3·151)}$$

The left-hand side of (3·15) is $O(e^{-c}) = O(ce^{-c})$ if $c \geqq 1$. We may therefore suppose $c<1$. In (3·14) we may write

$$\frac{e^{-\frac{1}{2}c}}{1-e^{-c}} = \frac{1}{c} - \frac{c}{24} + O(c^3), \quad \text{.........(3·16)}$$

$$e^{\beta_n c} - 1 = \beta_n c + \tfrac{1}{2}\beta_n^2 c^2 + O(c^3). \quad \text{.........(3·17)}$$

Since $c_1>c$ and $|\beta_n|<1$, we have

$$c^p \sum \beta_n^q e^{-nc_1} = O(c^{p-1}e^{-c})$$

for all positive integral values of $p$ and $q$. Hence the left-hand side of (3·14) takes the form

$$\sum_1^\infty \alpha_m e^{-mc} + \sum_1^\infty \beta_n e^{-nc_1} + \tfrac{1}{2}c\sum_1^\infty \beta_n^2 e^{-nc_1} + O(ce^{-c}). \quad \text{.........(3·18)}$$

Also, by (3·16),

$$\frac{e^{-\frac{1}{2}c}}{1-e^{-c}}\frac{e^{-c_1}}{1-e^{-c_1}}=\left(\frac{1}{c}-\frac{c}{24}\right)\frac{e^{-c_1}}{1-e^{-c_1}}+O(ce^{-c})$$

$$=\frac{1}{c}\frac{e^{-c_1}}{1-e^{-c_1}}-\frac{c}{24}\sum_1^{\infty}e^{-nc_1}+O(ce^{-c}). \qquad (3·19)$$

Hence (3·14) takes the form (3·15).

3·2. *Lemma E. If $\sigma$ is sufficiently large,*

$$\frac{1}{\Gamma(s)}\int_0^{\infty}c^{s-1}v(c,\theta)\,dc=\theta\zeta(s-1)-\tfrac{1}{2}\zeta(s)-\frac{\theta^{s-1}\zeta(s-1)}{s-1}. \qquad (3·21)$$

*This function is an integral function of s.*

The equation (3·21) follows at once from (3·151) by direct integration. It may be verified at once that the right-hand side is regular at its only possible singularities, viz. $s=1$ and $s=2$.

3·3. In what follows we denote by $\mathbf{D}(\alpha)$ a finite domain in the plane of $s$, all of whose points satisfy $\sigma \geqq \alpha+\delta>\alpha$; and by $R(s,\alpha)$ a function regular and bounded in $\mathbf{D}(\alpha)$. It is to be understood that the upper bound of such a function depends upon the form of $\mathbf{D}$, and in particular upon $\delta$, *but not upon* $\theta$, and that the $O$'s which we use are also uniform with respect to $\theta$.

*Lemma F. If $\psi(c,\theta)=O(c^q e^{-c})$, where $q \geqq 0$, then*

$$\chi(s,\theta)=\frac{1}{\Gamma(s)}\int_0^{\infty}\psi(c,\theta)\,c^{s-1}dc=R(s,-q). \qquad (3·31)$$

For the integral is uniformly convergent in $\mathbf{D}(-q)$.

*Lemma G. The function*

$$F_1(s,\theta)+\theta^s F_1(s,\theta_1)+\tfrac{1}{2}s\theta^{s+1}F_2(s+1,\theta_1) \qquad (3·32)$$

*is regular for $\sigma>-1$.*

Supposing first $\sigma$ sufficiently large, multiply (3·15) by $c^{s-1}/\Gamma(s)$, and integrate from $c=0$ to $c=\infty$. The result then follows immediately from Lemmas E and F. We obtain in fact

$$F_1(s,\theta)+\theta^s F_1(s,\theta_1)+\tfrac{1}{2}s\theta^{s+1}F_2(s+1,\theta_1)=\theta\zeta(s-1)-\tfrac{1}{2}\zeta(s)-\frac{\theta^{s-1}\zeta(s-1)}{s-1}+R(s,-1), \qquad (3·33)$$

and

$$Z_1(s)=Z_1(s,\theta)=\theta\zeta(s-1)-\tfrac{1}{2}\zeta(s)-\frac{\theta^{s-1}\zeta(s-1)}{s-1} \qquad (3·34)$$

is an integral function.

As a corollary, we have

THEOREM 2. *The function $F_1(s,\theta)+\theta^s F_1(s,\theta_1)$ is regular for $\sigma>0$.*

For $F_2(s+1,\theta_1)$ is plainly a function $R(s,0)$.

3·4. We have, from (3·33),

$$F_1(s,\theta)+\theta^s F_1(s,\theta_1)=Z_1(s,\theta)+R(s,0). \qquad (3·41)$$

Similarly

$$F_1(s,\theta_1)+\theta_1^s F_1(s,\theta_2)=Z_1(s,\theta_1)+R(s,0), \qquad (3·411)$$

and so on generally. From the first $n$ such equations we deduce

$$F_1(s,\theta)+(-1)^{n-1}(\theta\theta_1\ldots\theta_{n-1})^s F_1(s,\theta_n)=\Phi_n+\Psi_n, \qquad (3·42)$$

where
$$\Phi_n = \sum_{\nu=0}^{n-1} (-1)^\nu (\theta\theta_1 \dots \theta_{\nu-1})^s Z_1(s, \theta_\nu), \quad \text{.......................(3·421)}$$
$$\Psi_n = \sum_{\nu=0}^{n-1} (-1)^\nu (\theta\theta_1 \dots \theta_{\nu-1})^s R(s, 0)^*. \quad \text{.......................(3·422)}$$

We suppose for the moment that $\sigma > 2$.

Then
$$|F_1(s, \theta_n)| < A$$
where $A$ is independent of $n$ and $s$, and the second term on the left-hand side of (3·42) tends to zero. Similarly the functions $\Phi_n$ and $\Psi_n$ tend to
$$\Phi(s) = \sum_{\nu=0}^{\infty} (-1)^\nu (\theta\theta_1 \dots \theta_{\nu-1})^s Z_1(s, \theta_\nu), \quad \text{.......................(3·431)}$$
$$\Psi(s) = \sum_{\nu=0}^{\infty} (-1)^\nu (\theta\theta_1 \dots \theta_{\nu-1})^s R(s, 0), \quad \text{.......................(3·432)}$$
respectively; and
$$F_1(s, \theta) = \Phi(s) + \Psi(s), \quad \text{.................................(3·44)}$$
for $\sigma > 2$. This relation between analytic functions holds throughout any region in which each of them is regular. The function $\Psi(s)$ is plainly regular for $\sigma > 0$, since $\theta_\nu \theta_{\nu+1} < \frac{1}{2}$. We thus obtain

THEOREM 3. *The function*
$$F_1(s, \theta) - \Phi(s) \quad \text{.......................................(3·45)}$$
*is regular for* $\sigma > 0$.

The study of the singularities of $F_1(s, \theta)$, for $\sigma > 0$, is thus reduced to that of the singularities of $\Phi(s)$ in the same region.

3·5. THEOREM 4. *If $\theta$ is of class $\lambda$, then each of the functions*
$$F_1(s, \theta), \quad \Phi(s)$$
*is regular for*
$$\sigma > \sigma_1 = 1 - \frac{1}{1+\lambda} = \frac{\lambda}{1+\lambda}.$$
*If $\lambda > 0$, then the line $\sigma = \sigma_1$ is a singular line for each function.*

We observe first that
$$Z_1(s) = O(\theta^{\sigma-1}) + O(1), \quad \text{................................(3·51)}$$
uniformly in $\theta$, on any closed curve $C$ which does not pass through either of the points $s = 1$ or $s = 2$. The series for $\Phi(s)$ is thus the sum of two series, of the types
$$\Sigma\, O(\theta\theta_1 \dots \theta_{\nu-1})^\sigma, \quad \Sigma\, O(\theta\theta_1 \dots \theta_{\nu-1})^\sigma \theta_\nu^{\sigma-1}$$
respectively. The first series is uniformly convergent on $C$ if $C$ lies in any half-plane $\sigma \geqq \delta > 0$. The second is convergent if
$$\sigma + (\sigma - 1)\lambda > 0,$$
*i.e.* if $\sigma > \sigma_1$, and is uniformly convergent on $C$ if $C$ lies in any half-plane $\sigma \geqq \sigma_1 + \delta > \sigma_1$. It follows that $\Phi(s)$, and therefore $F_1(s, \theta)$, is regular inside any curve $C$ subject to these conditions, and therefore for $\sigma > \sigma_1$.

It remains to show that, when $\lambda > 0$, $\sigma = \sigma_1$ is a singular line, and it is plainly enough, after what precedes, to show that the line is singular for
$$\Phi_1(s) = \frac{\zeta(s-1)}{s-1} \Sigma (-1)^\nu (\theta\theta_1 \dots \theta_{\nu-1})^s \theta_\nu^{s-1} = \frac{\zeta(s-1)}{s-1} X(s), \quad \text{...........(3·52)}$$
or for $X(s)$. We choose $\delta > 0$, and divide the $\nu$'s into two classes $\nu'$, $\nu''$, writing $\nu = \nu'$ if
$$\theta_\nu < (\theta\theta_1 \dots \theta_{\nu-1})^{\lambda-\delta} \quad \text{.....................................(3·53)}$$

* $R(s, 0)$ is of course a different function in different terms of this series.

and $\nu=\nu''$ in the contrary case. In virtue of the definition of $\lambda$, there are, for every $\delta$, an infinity of $\nu'$'s.

We write
$$X(s)=\sum_{\nu}=\sum_{\nu'}+\sum_{\nu''}=X'(s)+X''(s). \quad \text{.........................(3·54)}$$

The series for $X''$ is absolutely convergent if $\sigma+(\sigma-1)(\lambda-\delta)>0$ or
$$\sigma>\frac{\lambda-\delta}{1+\lambda-\delta},$$
and the number on the right-hand side is less than $\sigma_1$. Hence $X''$ is regular across the line $\sigma=\sigma_1$. It is therefore sufficient to prove the line singular for $X'$.

Suppose that the values of $\nu'$ are $\nu_1, \nu_2, \ldots, \nu_k, \ldots$, and write
$$e^{-\lambda_k}=\theta\theta_1\ldots\theta_{\nu_k}. \quad \text{.....................................(3·55)}$$

Then the series for $X'(s)$, viz.
$$\sum\frac{(-1)^{\nu_k}}{\theta_{\nu_k}}(\theta\theta_1\ldots\theta_{\nu_k})^s=\sum\frac{(-1)^{\nu_k}}{\theta_{\nu_k}}e^{-\lambda_k s}$$
is a Dirichlet's series of the type $\sum a_k e^{-\lambda_k s}$, and
$$\lambda_{k+1}-\lambda_k=\log\frac{1}{\theta_{\nu_k+1}\theta_{\nu_k+2}\cdots\theta_{\nu_{k+1}}}\geqslant\log\frac{1}{\theta_{\nu_{k+1}}}>(\lambda-\delta)\log\frac{1}{\theta\theta_1\ldots\theta_{\nu_{k+1}-1}}\to\infty$$
when $k\to\infty$. It follows, by a theorem of Wennberg*, that the line $\sigma=\sigma_1$ is singular for $X'$, which completes the proof of the theorem.

We have supposed $0<\lambda<\infty$. When $\lambda=\infty$ the result is still valid, $\sigma=1$ being a singular line; and only trivial modifications are needed in the proof. The case $\lambda=0$ is much more difficult. It appears to be true that $\sigma=0$ is then a singular line, except in the special case in which $\theta$ is a quadratic surd; but we are unable to prove this rigorously. The exceptional case is that studied by Hecke.

3·6. Suppose in particular that $\theta$ is a quadratic surd. The continued fraction for $\theta$ is then periodic, and we have
$$\theta_{r+km}=\theta_r \quad (r\geqq\rho,\ k=1,2,\ldots),$$
if $\rho$ is the number of non-repeated $\theta$'s and $m$ the length of the period.

In this case $F_2(s,\theta)$ is, by Theorem 1, regular for $\sigma>-1$. It follows from Lemma G that each of the functions
$$F_1(s,\theta)+(-1)^{\rho-1}(\theta\theta_1\ldots\theta_{\rho-1})^s F_1(s,\theta_\rho),$$
$$F_1(s,\theta_\rho)+(-1)^{m-1}(\theta_\rho\theta_{\rho+1}\ldots\theta_{\rho+m-1})^s F_1(s,\theta_{\rho+m})$$
is regular for $\sigma>-1$. But the last function is
$$(1+(-1)^{m-1}\Theta^s)F_1(s,\theta_\rho).$$
It follows that $F_1(s,\theta_\rho)$, and therefore $F_1(s,\theta)$, is regular for $\sigma>-1$, except possibly where
$$1+(-1)^{m-1}\Theta^s=0,$$
at which points it may have simple poles. These points are the points
$$s=\frac{k\pi i}{\log\Theta}$$
where $k$ is an arbitrary odd or arbitrary even integer, according as $m$ is odd or even.

* Wennberg, 'Zur Theorie der Dirichlet'schen Reihen', *Inaugural dissertation*, Upsala, 1920, 3—7. It has been shown by Carlson and Landau that the result is true under the more general conditions
$$\lambda_{n+1}-\lambda_n>A,\quad \lambda_n/n\to\infty.$$
See F. Carlson and E. Landau, 'Neuer Beweis und Verallgemeinerungen des Fabryschen Lückensatzes', *Göttinger Nachrichten*, 1921, 184—188.

3·7. There appears to be no doubt of the truth of the following propositions:

($ak$) *$F_k(s)$ is regular for $\sigma > \sigma_k$;*

($bk$) *$\sigma = \sigma_k$ is a singular line for $F_k(s)$ whenever $\lambda > 0$;*

($ck$) *$\sigma = \sigma_k$ is singular even when $\lambda = 0$, except when $\theta$ is quadratic;*

($dk$) *$F_k(s)$ is meromorphic when $\theta$ is quadratic; its poles are all simple; and they are situated at some or all of a doubly infinite system of points distributed at equal distances along the lines*

$$\sigma = 1 - k - 2p \quad (p = 0, 1, \ldots);$$

($ek$) *$F_k(s, \theta) + (-1)^{k-1}\theta^{s+k-1}F_k(s, \theta_1)$ is regular for $\sigma > \sigma_{k+1} - 1$; and $\sigma = \sigma_{k+1} - 1$ is a singular line for the function when $\lambda > 0$:*

and a complete theory of the functions would contain proofs of these propositions in full generality.

Of these propositions we have proved ($ak$), in § 2 when $k > 1$ and in § 3·5 when $k = 1$.

We are unable to prove ($ck$) in any case. The case in which $\theta$ is quadratic is doubtless best treated by the deeper methods of Hecke. We have however shown, in § 3·6, that our method will accomplish *something* in the direction indicated by ($dk$).

There remain the propositions ($bk$) and ($ek$), of which, at present, we have proved ($b\,1$) only. We proceed now to the general proof. The particular case contains most of the leading ideas, and we have condensed the general argument wherever the ground is familiar. In what follows the $A$'s, $O$'s, and $R(s, \alpha)$'s depend on $k$ in addition to the regions $\mathbf{D}$; they are either independent of $\theta$, as in § 3·3, or at any rate, when we have to consider a sequence of irrationals $\theta, \theta_1, \theta_2, \ldots$, of the $n$ in $\theta_n$.

*Lemma H. If $k > 1$ we have, throughout $\mathbf{D}(\sigma_k)$,*

$$|F_k(s, \theta)| < A\Sigma\nu^{-k}|\overline{\nu\theta}|^{\sigma-1-\frac{1}{2}\delta} + A = AT + A \leq A\Sigma\nu^{-k}|\overline{\nu\theta}|^{\sigma_k-1+\frac{1}{2}\delta} + A.$$

This is a straightforward deduction from the results of §§ 2·1, 2·2. By (2·231) and (2·232),

$$|G_k(s)| \leq \Sigma|X(\nu)|\nu^{-k} \leq \Sigma(|X_0(\nu)| + A|\mathbf{x}_m|^{\sigma-1})\nu^{-k}.$$

Also $|X_0(\nu)| < A$, by (2·22); and, since $A|\overline{\nu\theta}| < |\mathbf{x}_m| < A$, we have

$$|\mathbf{x}_m|^{\sigma-1} < A + A|\overline{\nu\theta}|^{\sigma-1}.$$

Hence $$|G_k(s)| < A\Sigma\nu^{-k}|\overline{\nu\theta}|^{\sigma-1} + A < AT + A. \quad \text{.......................(3·71)}$$

Let $\mathbf{D}'$ be the domain obtained by removing from $\mathbf{D}$ circles $C_1, C_2, \ldots$ of radius $\frac{1}{4}\delta$ surrounding such poles $1, 2, \ldots$ of $\Gamma(1-s)$ as fall in $\mathbf{D}$. Then

$$|g_k(s)| = |\Gamma(1-s)||G_k(s)| < A\Sigma\nu^{-k}|\overline{\nu\theta}|^{\sigma-1} + A < AT + A \quad \text{...........(3·72)}$$

in $\mathbf{D}'$. On $C_1$, $1 - \frac{1}{4}\delta \leq \sigma \leq 1 + \frac{1}{4}\delta$, and it is easily deduced that

$$|g_k(s)| < A\Sigma\nu^{-k}|\overline{\nu\theta}|^{-\frac{1}{2}\delta} + A < AT + A \quad \text{.......................(3·73)}$$

on $C_1$. The middle term here is independent of $\sigma$, and $g_k(s)$ is regular for $s = 1$, so that the inequalities (3·73) are valid also inside $C_1$. Similarly it may be shown that $|g_k(s)| < AT + A$ throughout $C_2, C_3, \ldots$, and so, by (3·72), throughout $\mathbf{D}$. A similar argument may be applied to $h_k(s)$, and the lemma follows, since $F_k(s)$ is a linear combination of the two functions.

*Lemma K. Throughout $\mathbf{D}(\sigma_{k+1})$*

$$|F_{k+r}(s + r, \theta_n)| < A + A(\theta\theta_1 \ldots \theta_{n-1})^{(\sigma-\frac{1}{2}\delta)(\lambda+\delta)} \quad (1 \leq r \leq k)^*.$$

* The important case is $\sigma < 0$. When $\sigma \geqq \delta$ the second term may be absorbed in the first; the proof will be clearer if $\sigma < 0$ is thought of as the standard case.

The left-hand side is less than

$$A\Sigma\nu^{-k-r}\,|\,\overline{\nu\theta_n}\,|^{\sigma+r-1-\frac{1}{2}\delta}+A<A\Sigma\nu^{-k-1}\,|\,\overline{\nu\theta_n}\,|^{\sigma-\frac{1}{2}\delta}+A,$$

by Lemma H. Let $\epsilon=\frac{1}{2}\delta/(k+1)<\delta$, and $h=\lambda+1+\epsilon$. Then

$$\frac{1}{\theta_{n+l}}<\frac{A}{(\theta\theta_1\ldots\theta_{n+l-1})^{h-1}}<\frac{At_n^{-1}}{(\theta_n\ldots\theta_{n+l-1})^{h-1}},\qquad(3{\cdot}74)$$

where $$t_n=(\theta\theta_1\ldots\theta_{n-1})^{h-1}=(\theta\ldots\theta_{n-1})^{\lambda+\epsilon}.\qquad(3{\cdot}75)$$

If now $P_m/Q_m$ is the $m$th convergent of the continued fraction for $\theta_n$, (3·74) implies that $Q_m<At_n^{-1}Q^h_{m-1}$, by Lemma 2 of our paper (4), and therefore that

$$|\,\nu\theta_n\,|>At_n\nu^{-h}.$$

Hence

$$F_{k+r}(s+r,\theta_n)<\sum_\nu\nu^{-k-1}\{A+(At_n\nu^{-h})^{\sigma-\frac{1}{2}\delta}\}<A+At_n^{\sigma-\frac{1}{2}\delta}\sum_\nu\nu^{-k-1-h(\sigma-\frac{1}{2}\delta)}\leqq A+At_n^{\sigma-\frac{1}{2}\delta}\sum_\nu\nu^{-k-1-h(\sigma_{k+1}+\frac{1}{2}\delta)}.$$

The index of $\nu$ is

$$-k-1-(1+\lambda+\epsilon)\left(1-\frac{k+1}{1+\lambda}+\tfrac{1}{2}\delta\right)=-1-\epsilon-\left(\tfrac{1}{2}\delta-\frac{(k+1)\,\epsilon}{1+\lambda}\right)-\lambda-\tfrac{1}{2}\delta\lambda-\delta\epsilon<-1-\epsilon,$$

so that the series last written is convergent. Hence

$$|\,F_{k+r}(s+r,\theta_n)\,|<A+A\,(\theta\ldots\theta_{n-1})^{(\sigma-\frac{1}{2}\delta)(\lambda+\epsilon)}<A+A\,(\theta\ldots\theta_{n-1})^{(\sigma-\frac{1}{2}\delta)(\lambda+\delta)},$$

the result of the lemma.

3·8. We return now to the identity (3·13), and we equate the coefficients of $\xi^{k-1}/k!$ in the Laurent expansions of the two sides. If we define $\phi_0(x)$ to be unity, then

$$\frac{e^{xz}}{e^z-1}=\sum_{r=0}^{\infty}\frac{\phi_r(x)}{r!}z^{r-1}.$$

Hence the coefficient in the first term on the left is

$$\Sigma\{\phi_k(\alpha_m+\tfrac{1}{2})-\phi_k(\tfrac{1}{2})\}\,e^{-mc}=\Sigma\phi_k(m\theta)\,e^{-mc}-\phi_k(\tfrac{1}{2})\,\Sigma e^{-mc}.\qquad(3{\cdot}81)$$

The coefficient of $\xi^{k-1}/k!$ in the second term is

$$k\,(-\theta)^{k-1}\Sigma\,u_ne^{-nc_1},\qquad(3{\cdot}82)$$

where $$u_n=u_n(c)=\left(\frac{d}{dc}\right)^{k-1}\left(\frac{e^{-\frac{1}{2}c}}{1-e^{-c}}(e^{\beta_nc}-1)\right)=\left(\frac{d}{dc}\right)^{k-1}\left(\frac{e^{\omega c}-e^{\frac{1}{2}c}}{e^c-1}\right),\qquad(3{\cdot}821)$$

$$\omega=\omega_n=\beta_n+\tfrac{1}{2}=(n\theta_1).\qquad(3{\cdot}822)$$

Now $$u_n(c)=\sum_{r=0}^{k+1}\frac{u_n^{(r)}(0)}{r!}c^r+\frac{u_n^{(k+2)}(\vartheta c)}{(k+2)!}c^{k+2}\quad(0<\vartheta<1)$$

$$=\sum_{r=0}^{k+1}\left[\left(\frac{d}{dc}\right)^{k+r-1}\left(\frac{e^{\omega c}-e^{\frac{1}{2}c}}{e^c-1}\right)\right]_0\frac{c^r}{r!}+\left[\left(\frac{d}{dc}\right)^{2k+1}\left(\frac{e^{\omega c}-e^{\frac{1}{2}c}}{e^c-1}\right)\right]_{\vartheta c}\frac{c^{k+2}}{(k+2)!}$$

$$=\sum_{r=0}^{k+1}\frac{1}{k+r}\{\phi_{k+r}(\omega)-\phi_{k+r}(\tfrac{1}{2})\}\frac{c^r}{r!}+\Phi(\vartheta c)\frac{c^{k+2}}{(k+2)!},\qquad(3{\cdot}83)$$

say; and it is easily verified that

$$|\,\Phi(x)\,|<A\quad(x>0).\qquad(3{\cdot}84)$$

Summing up from (3·81), (3·82), (3·83), and (3·84), we find that the coefficient of $\xi^{k-1}/k!$ in the left-hand side of (3·13) is

$$\Sigma\phi_k(m\theta)\,e^{-mc}+(-\theta)^{k-1}\,\Sigma\,\phi_k(n\theta_1)\,e^{-nc_1}+(-\theta)^{k-1}\sum_{r=1}^{k+1}\frac{k}{k+r}\frac{c^r}{r!}\Sigma\,\phi_{k+r}(n\theta_1)\,e^{-nc_1}$$

$$-\phi_k(\tfrac{1}{2})\,\Sigma e^{-mc}-(-\theta)^{k-1}\sum_{r=0}^{k+1}\frac{k}{k+r}\phi_{k+r}(\tfrac{1}{2})\frac{c^r}{r!}\Sigma\,e^{-nc_1}+O(c^{k+2})\,\Sigma e^{-nc_1}.\qquad(3{\cdot}85)$$

We consider next the coefficient of $\xi^{k-1}/k!$ in $W$, the right-hand side of (3·13). Now $W$ is regular at $\xi = 0$, and, expanding formally, we have

$$W = \left\{ \sum_{r=0}^{\infty} \frac{\phi_r(0)}{r!} \xi^{r-1} \right\} \left\{ \sum_{r=0}^{\infty} \frac{(-\theta\xi)^r}{r!} \left(\frac{d}{dc}\right)^r \left(\frac{e^{-c}}{1-e^{-c}}\right) \right\}$$
$$- \frac{e^{-c}}{1-e^{-c}} \left\{ \sum_{r=0}^{\infty} \frac{\phi_r(\frac{1}{2})}{r!} \xi^{r-1} \right\} - \frac{e^{-c_1}}{1-e^{-c_1}} \left\{ \sum_{r=0}^{\infty} \frac{(-\theta\xi)^r}{r!} \left(\frac{d}{dc}\right)^r \left(\frac{e^{-\frac{1}{2}c}}{1-e^{-c}}\right) \right\}.$$

Collecting the coefficient of $\xi^{k-1}/k!$, and equating it to (3·85), we obtain

$$\Sigma \phi_k(m\theta) e^{-mc} + (-\theta)^{k-1} \Sigma \phi_k(n\theta_1) e^{-nc_1} + \sum_{r=1}^{k+1} (-\theta)^{k-1} \frac{k}{k+r} \frac{c^r}{r!} \Sigma \phi_{k+r}(n\theta_1) e^{-nc_1}$$
$$= O(c^{k+2}) \frac{1}{e^{c_1}-1} + V_k(c), \quad \text{......(3·861)}$$

where

$$V_k = \sum_{r=0}^{k} \binom{k}{r} \phi_r(0) (-\theta)^{k-r} \left(\frac{d}{dc}\right)^{k-r} \left(\frac{1}{e^c-1}\right) - \frac{k(-\theta)^{k-1}}{e^{c_1}-1} \left\{ \left(\frac{d}{dc}\right)^{k-1} \left(\frac{e^{\frac{1}{2}c}}{e^c-1}\right) - \sum_{r=0}^{k-1} \frac{\phi_{k-r}(\frac{1}{2})}{k+r} \frac{c^r}{r!} \right\}.$$
$$\text{.........(3·862)}$$

In (3·862) we associate a term $-\dfrac{1}{c}$ with $\dfrac{1}{e^c-1}$ and $\dfrac{e^{\frac{1}{2}c}}{e^c-1}$, perform some trivial rearrangements and reductions*, and obtain the alternative expressions

$$V_k = \sum_{r=0}^{k} \binom{k}{r} \phi_r(0) (-\theta)^{k-r} \left(\frac{d}{dc}\right)^{k-r} \left(\frac{1}{e^c-1} - \frac{1}{c}\right) - \frac{k!}{\theta} \left(\frac{c}{\theta}\right)^{-k} \left\{ \frac{1}{e^{c_1}-1} - \sum_{r=0}^{k} \frac{\phi_r(0)}{r!} \left(\frac{c}{\theta}\right)^{r-1} \right\}$$
$$+ \frac{(-1)^k k \theta^{k-1}}{e^{c_1}-1} \left\{ \left(\frac{d}{dc}\right)^{k-1} \left(\frac{e^{\frac{1}{2}c}}{e^c-1} - \frac{1}{c}\right) - \sum_{r=0}^{k} \frac{\phi_{k+r}(\frac{1}{2})}{k+r} \frac{c^r}{r!} \right\} \quad \text{.........(3·863)}$$

or

$$V_k = V_{k,1} + V_{k,2} + O(c^{k+2}) \frac{1}{e^{c_1}-1}, \quad \text{..........................(3·864)}$$

where $V_{k,1}$ and $V_{k,2}$ are the first and second terms on the right-hand side of (3·863)†.

We now multiply (3·861) by $c^{s-1}/\Gamma(s)$, integrate from $c = 0$ to $c = \infty$, and obtain

$$F_k(s, \theta) + (-1)^{k-1} \theta^{k-1+s} F_k(s, \theta_1) + (-1)^{k-1} \sum_{r=1}^{k+1} \theta^{k-1+s+r} \frac{k}{k+r} \frac{\Gamma(s+r)}{\Gamma(s)\Gamma(r+1)} F_{k+r}(s+r, \theta_1)$$
$$= R(s, -k-1) + \frac{1}{\Gamma(s)} \int_0^{\infty} \{V_{k,1}(c) + V_{k,2}(c)\} c^{s-1} dc$$
$$= R(s, -k-1) + Z_k(s, \theta), \quad \text{...............................................(3·871)}$$

say. By (3·864) and (3·862),

$$V_{k,1} + V_{k,2} = V_k + O(c^{k+2} e^{-c}) = O(c^{k+2} e^{-c}) \quad (c > 1).$$

Further, the formulae which define $V_{k,1}$ and $V_{k,2}$ show that these functions tend to limits as $c \to 0$. It follows that

$$\int_0^{\infty} (V_{k,1} + V_{k,2}) c^{s-1} dc$$

exists, and defines a function of $s$ regular for $\sigma > 0$.

Next, we observe that

$$\frac{1}{\Gamma(s)} \int_0^{\infty} \left(\frac{d}{dc}\right)^{k-r} \left(\frac{1}{e^c-1} - \frac{1}{c}\right) c^{s-1} dc = \eta_r(s),$$

and

$$\frac{1}{\Gamma(s)} \int_0^{\infty} V_{k,2}(c) c^{s-1} dc = -\frac{k!\,\theta^{s-1}}{\Gamma(s)} \int_0^{\infty} \left\{ \frac{1}{e^{c_1}-1} - \sum_{r=0}^{k} \frac{\phi_r(0)}{r!} c_1^{r-1} \right\} c_1^{-k+s-1} dc_1 = \theta^{s-1} z(s)$$

* Observing in particular that the term for which $r = k+1$ vanishes, since $\phi_{2k+1}(\frac{1}{2}) = 0$.

† The expression in curly brackets in the third term is $O(c^{k+2})$.

exist and define functions regular for $0<\sigma<1$. Hence

$$Z_k(s,\theta)=\sum_{r=0}^{k} A_r\theta^{k-r}\eta_r(s)+\theta^{s-1}z(s), \qquad (3\cdot88)$$

this equation giving, moreover, the analytic continuation of $Z_k$ wherever $\eta_r$ and z are regular. Now $\eta_r$ and z do not contain $\theta$; they belong to a well-known class of integrals; and it may be shown that they are regular everywhere, except possibly for simple poles at certain *positive* integral values of $s$*. This being so, and if we suppose positive integral values of $s$ to be excluded from $\mathbf{D}(-k-1)$ by circles of radius $\delta$, we have

$$\sum_{r=0}^{k} A_r\theta^{k-r}\eta_r(s)=R(s,-k-1);$$

for only positive powers of $\theta$ occur on the left-hand side. Hence, from (3·871),

$$F_k(s,\theta)+(-1)^{k-1}\theta^{k-1+s}F_k(s,\theta_1)=R(s,-k-1)+\theta^{s-1}z(s)+Q(s,\theta_1), \qquad (3\cdot89)$$

where $$Q(s,\theta_1)=(-1)^k\sum_{r=1}^{k+1}\theta^{k-1+s+r}\frac{k}{k+r}\frac{\Gamma(s+r)}{\Gamma(s)\Gamma(r+1)}F_{k+r}(s+r,\theta_1). \qquad (3\cdot891)$$

3·9. We are now in a position to prove that $\sigma=\sigma_k$ is a barrier for $F_k(s,\theta)$ when $\lambda>0$. The case $\lambda=\infty$ is comparatively trivial, and we suppose that $0<\lambda<\infty$. Then

$$\sigma_k>1-k, \quad \sigma_k>\sigma_{k+1}.$$

We write $\theta_{n-1}$ for $\theta$ in (3·89), multiply by $(-1)^{kn}(\theta\theta_1\ldots\theta_{n-2})^{k-1+s}$, and sum as in § 3·4. We then apply our former argument, which shows on the one hand that

$$\Sigma(-1)^{kn}(\theta\ldots\theta_{n-2})^{k-1+s}R(s,-k-1)$$

is regular for $\sigma>1-k$ (except possibly for certain positive integral values of $s$), and therefore regular across $\sigma=\sigma_k$; and on the other that

$$\Sigma(-1)^{kn}(\theta\ldots\theta_{n-2})^{k-1+s}\theta_{n-1}^{s-1}z(s)$$

has $\sigma=\sigma_k$ for a barrier. To complete the proof of $(bk)$ it is sufficient to show that

$$\Sigma(-1)^{kn}(\theta\ldots\theta_{n-2})^{k-1+s}Q(s,\theta_n)$$

is regular across $\sigma=\sigma_k$; and this is true provided that, for some $\sigma'<\sigma_k$, the series

$$\Sigma(\theta\ldots\theta_{n-2})^{k-1+\sigma}\ Q(s,\theta_n) \qquad (3\cdot91)$$

is uniformly convergent in the part of $D(\sigma_{k+1})$ for which $\sigma\geqq\sigma'$. Now every term in (3·891) contains $\theta$ to the power $k-1+\sigma$ at least; hence, by Lemma K,

$$|Q(s,\theta_n)|<A\theta_{n-1}^{k-1+\sigma}\{1+(\theta\ldots\theta_{n-1})^{(\sigma-\frac{1}{2}\delta)(\lambda+\delta)}\},$$

and the general term in (3·91) is less than

$$A(\theta\ldots\theta_{n-1})^{k-1+\sigma'}+A(\theta\ldots\theta_{n-1})^{k-1+\sigma'+(\sigma'-\frac{1}{2}\delta)(\lambda+\delta)}.$$

The first index is positive if $\sigma_k-\sigma'$ is small enough, since $\sigma_k>1-k$; and the second is

$$k-1+\sigma_k+\sigma_k(\lambda+\delta)-(\sigma_k-\sigma')(1+\lambda+\delta)-\tfrac{1}{2}\delta(\lambda+\delta)=\lambda+\delta\sigma_k-(\sigma_k-\sigma')(1+\lambda+\delta)-\tfrac{1}{2}\delta(\lambda+\delta),$$

and is positive if $\delta$ and $\sigma_k-\sigma'$ are small enough. This establishes the uniform convergence of the series (3·91), and so finally the general result $(bk)$.

There remains $(ek)$; and for the proof of this the material we have already is sufficient. We observe that

$$\sigma_{k+1}-1\geqq-k-1, \quad \sigma_{k+r}\leqq\sigma_{k+1} \qquad (r\geqq1).$$

* See for example A. Hurwitz, 'Ueber die Anwendung eines functionentheoretischen Principes auf gewisse bestimmte Integrale', *Math. Annalen*, 53 (1900), 220—224.

From these facts, and from (3·871) and (3·88), it follows that the only possible singularities of

$$G_k(s) = F_k(s, \theta) + (-1)^{k-1}\theta^{s+k} F_k(s, \theta_1),$$

in $\sigma > \sigma_{k+1} - 1$, are the singularities of $Z_k$ in this region. These can occur only for positive integral values of $s$. On the other hand (3·871) shows that (when $\lambda < \infty$)* $Z_k$ is regular in $\sigma \geqq \sigma'$, where $\sigma' < 1$†. Hence $Z_k$ and $G_k$ are regular in $\sigma > \sigma_{k+1} - 1$.

On the other hand, if $\lambda > 0$,

$$\sigma_{k+1} - 1 > -k,$$

and $\sigma_{k+r}$ is a strictly decreasing function of $r$. It follows from (3·871) that $G_k$ and

$$(-1)^k \theta^{s+k} \frac{k}{k+1} s F_{k-1}(s+1, \theta_1)$$

are equi-singular in the region $\sigma > \text{Max}(-k-1, \sigma_{k+2} - 1)$. Since $F_{k-1}(s+1, \theta_1)$ has a barrier $\sigma = \sigma_{k+1} - 1$ in this region, this line is also a barrier for $G_k$.

We have therefore proved

THEOREM 5. *If $\lambda > 0$, the line $\sigma = \sigma_k$ is a singular line of $F_k(s, \theta)$.*

THEOREM 6. *If $\lambda > 0$ and $\theta = 1/(a_1 + \theta_1)$, then the function*

$$F_k(s, \theta) + (-1)^{k-1}\theta^{s+k-1} F_k(s, \theta_1)$$

*is regular for $\sigma > \sigma_{k+1} - 1$; and the line $\sigma = \sigma_{k+1} - 1$ is a singular line of the function.*

4·1. We conclude with a brief discussion of the problem of the convergence or summability of the series $\Sigma\phi_k(n\theta) n^{-s}$ in the region of existence of the corresponding function $F_k(s, \theta)$. It will be seen that our conclusions may be roughly expressed by saying that whatever *could* be true *is* true. A Dirichlet's series cannot be summable outside its half-plane of regularity, and it cannot be summable $(C, r)$ unless its $n$th term is of the form $o(n^r)$: we shall show that our series is summable (with least possible order) except when these restrictions apply.

THEOREM 7. *The series $\Sigma\phi_k(n\theta) n^{-s}$ is convergent if $\sigma > \sigma_k$, $\sigma > 0$; and summable $(C, -\sigma + \delta)$, for every positive $\delta$, if $\sigma > \sigma_k$, $\sigma < 0$.*

The case $\lambda = \infty$ is trivial, since $\sigma_k = 1$, and we suppose that $\lambda < \infty$. We may confine ourselves also to the case $k > 1$; for the result for $k = 1$ is an immediate deduction from the formula

$$\sum_{n<x} \alpha_n = O\left(x^{1 - \frac{1}{1+\lambda} + \epsilon}\right) = O(x^{\sigma_1 + \epsilon})\text{‡}. \qquad \text{(4·11)}$$

We shall in fact prove rather more than we have stated, when $k > 1$, viz. that the series is summable $(C, -\sigma' + \delta)$, where $\sigma' = \text{Min}(\sigma, 1)$.

When $k > 1$, $\phi_k(n\theta)$ is of the form $A(\psi_k(n\theta) \pm \psi_k(-n\theta))$, where $\psi_k(n\theta)$ is the function (2·121). It is therefore sufficient for our purpose to show that the series

$$\Sigma\psi_k(n\theta) n^{-s}, \quad \Sigma\psi_k(-n\theta) n^{-s},$$

are summable $(C, -\sigma' + \delta)$ for $\sigma > \sigma_k$. Further, it is enough to prove this for real values of $s$, and hence also, since $\psi_k(n\theta)$ and $\psi_k(-n\theta)$ are then conjugate imaginaries, enough to prove it for the first series. This we do by a series of lemmas.

* The case $\lambda = \infty$ is a trivial deduction from (3·89).

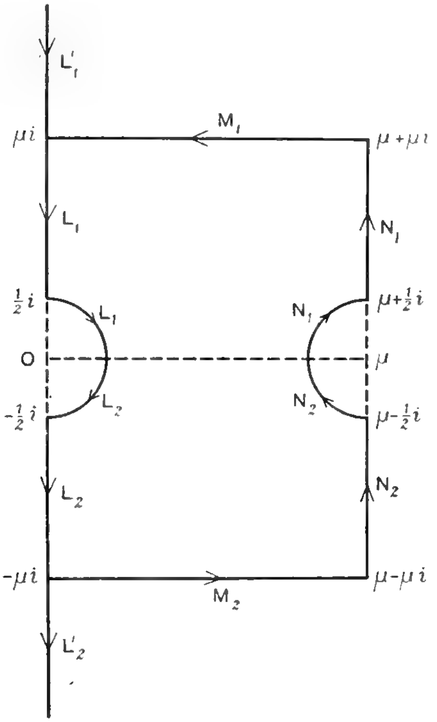
† Thus $Z_k$ is an integral function of $s$ when $\lambda < \infty$. This conclusion may be extended also to the case $\lambda = \infty$. For our argument shows that, in any case, $Z_k$ can have as a singularity at most a simple pole at $s = 1$, the residue being a rational function $r(\theta)$ of $\theta$. Since $r(\theta)$ vanishes for every $\theta$ for which $0 < \lambda < \infty$, it must be identically zero.

‡ See our paper (4), Theorem 2.

**4·2.** Let $\mu$ be a (large) positive integer, and let $0 < \phi < 1$, $-1 < \alpha < \beta \leq \alpha + 1$;

$$S(\mu, \phi) = \sum_{n<\mu} n^{\alpha} e^{2n\pi i\phi} \left(1 - \frac{n}{\mu}\right)^{\beta},$$

$$S(\phi) = \lim_{r \to 1-0} \Sigma n^{\alpha} e^{2n\pi i\phi} r^{n}.$$




Let $\Lambda_1$ be the contour $L_1 + L_1'$ of the figure (in which the sense of description is indicated by an arrow), $\Lambda_2$ be $L_2 + L_2'$, and $\Lambda$ be $\Lambda_1 + \Lambda_2$. Finally, let $C$ be the indented rectangle

$$L_1 + L_2 + M_2 + N_2 + N_1 + M_1.$$

In what follows $A$'s denote positive constants depending only on $k$, $\alpha$, and $\beta$, and the $O$'s have a corresponding meaning.

*Lemma L*: $|S(\phi)| < A\,\phi^{-(\alpha+1)}$.

This is a particular case of a known result. In fact the function $f(z)$ defined, for $|z| < 1$, by the series $\Sigma n^{\alpha} z^{n}$, has $z = 1$ for its sole singularity, and

$$|f(z)| < A\,|1 - z|^{-(\alpha+1)},$$

so that
$$|S(\phi)| = |f(e^{2\pi i\phi})| < A\,|1 - e^{2\pi i\phi}|^{-(\alpha+1)} < A\,\phi^{-(\alpha+1)}.$$

*Lemma M*: $S(\phi) = \displaystyle\int_{\Lambda_1+\Lambda_2} z^{\alpha} \frac{e^{2z\pi i\phi}}{e^{2z\pi i} - 1}\, dz.$

This is a particular case of a very general formula in the theory of residues*.

* See Lindelöf, *l.c.*, ch. v, § 53.

4·3. *Lemma N. If* $0 < |\phi| < \frac{1}{2}$, *then* $|S(\mu, \phi) - S(\phi)| < A|\phi|^{-(\beta+1)}\mu^{\alpha-\beta}$.

We may suppose that $0 < \phi < \frac{1}{2}$, and we begin by showing that

$$S(\mu, \phi) = \int_{L_1+L_2} z^{\alpha}\left(1-\frac{z}{\mu}\right)^{\beta}\frac{e^{2z\pi i\phi}}{e^{2z\pi i}-1}\,dz + O(T), \qquad (4·31)$$

$$S(\phi) = \int_{L_1+L_2} z^{\alpha}\frac{e^{2z\pi i\phi}}{e^{2z\pi i}-1}\,dz + O(T), \qquad (4·32)$$

where $T = \phi^{-(\beta+1)}\mu^{\alpha-\beta}$. In the first place, by the theorem of residues, we have

$$S(\mu, \phi) = \int_C z^{\alpha}\left(1-\frac{z}{\mu}\right)^{\beta}\frac{e^{2z\pi i\phi}}{e^{2z\pi i}-1}\,dz. \qquad (4·33)$$

Now

$$\left|\frac{e^{2z\pi i\phi}}{e^{2z\pi i}-1}\right| < \begin{cases} Ae^{-2\pi\phi y} & (y > A) \\ Ae^{2\pi(1-\phi)y} & (y < -A), \end{cases}$$

and so

$$\left|\frac{e^{2z\pi i\phi}}{e^{2z\pi i}-1}\right| < Ae^{-A\phi|y|} \quad (|y| > A).$$

Hence the contributions of $M_1$ and $M_2$ to the right-hand side of (4·33) are of the form

$$O(\mu^{\alpha+1}e^{-A\mu\phi}) = O(T)(\mu\phi)^{\beta+1}e^{-A\mu\phi} = O(T),$$

since $\beta + 1 > 0$, so that $(\mu\phi)^{\beta+1}e^{-A\mu\phi} < A$. Similarly the straight portions of $N_1$, $N_2$ contribute

$$O\left(\int_{\frac{1}{2}}^{\mu}|\mu + iy|^{\alpha}\left(\frac{y}{\mu}\right)^{\beta}e^{-Ay\phi}\,dy\right) = O(\mu^{\alpha-\beta})\int_{\frac{1}{2}}^{\mu}y^{\beta}e^{-Ay\phi}\,dy = O(\mu^{\alpha-\beta}\phi^{-(\beta+1)}) = O(T).$$

Lastly, the curved portions of $N_1$ and $N_2$ contribute $O(\mu^{\alpha-\beta}) = O(T)$. Thus the contour $C$, less $L_1 + L_2$, contributes $O(T)$, and (4·31) is proved.

For $S(\phi)$ we have

$$S(\phi) - \int_{L_1+L_2} z^{\alpha}\frac{e^{2z\pi i\phi}}{e^{2z\pi i}-1}\,dz = \int_{L_1'+L_2'} z^{\alpha}\frac{e^{2z\pi i\phi}}{e^{2z\pi i}-1}\,dz = O\left(\int_{\mu}^{\infty}y^{\alpha}e^{-Ay\phi}\,dy\right)$$

$$= O\left(e^{-A\mu\phi}\int_0^{\infty}y^{\alpha}e^{-Ay\phi}\,dy\right) = O\left(e^{-A\mu\phi}\phi^{-\alpha-1}\right) = O\left(e^{-A\mu\phi}(\mu\phi)^{\beta-\alpha}T\right) = O(T).$$

This is (4·32).

From (4·31) and (4·32) we deduce

$$S(\mu, \phi) - S(\phi) = \int_{L_1+L_2} z^{\alpha}\left\{\left(1-\frac{z}{\mu}\right)^{\beta} - 1\right\}\frac{e^{2z\pi i\phi}}{e^{2z\pi i}-1}\,dz + O(T).$$

Now

$$\left|\left(1-\frac{z}{\mu}\right)^{\beta} - 1\right| < A\left|\frac{z}{\mu}\right|$$

on $L_1 + L_2$. Hence the straight portions of $L_1 + L_2$ contribute

$$O\left(\int_{\frac{1}{2}}^{\mu}y^{\alpha}\cdot\frac{y}{\mu}\cdot e^{-Ay\phi}\,dy\right) = O\left(\int_{\frac{1}{2}}^{\mu}y^{\alpha}\cdot\left(\frac{y}{\mu}\right)^{\beta-\alpha}\cdot e^{-Ay\phi}\,dy\right) = O(\mu^{\alpha-\beta})\int_0^{\infty}y^{\beta}e^{-Ay\phi}\,dy = O(T),$$

since $\beta - \alpha \leqq 1$ and $y/\mu \leqq 1$. Finally the curved portions contribute

$$O\left(\frac{1}{\mu}\right) = O(\mu^{\alpha-\beta}) = O(T).$$

This completes the proof of Lemma N.

4·4. We can now deduce that $\Sigma\,\psi_k(n\theta)\,n^{-\sigma}$ is summable $(C, -\sigma' + \delta)$ for $\sigma > \sigma_k$. We may suppose that $\sigma < 1$, since a convergent series, whose general term is $O(1/n)$, is summable $(C, -1+\delta)$*. Also $\sigma_k < 1$, since $\lambda < \infty$. Hence we may suppose that $\sigma_k < \sigma = \sigma' < 1$. This being

* G. H. Hardy and J. E. Littlewood, 'Contributions to the arithmetic theory of series', *Proc. London Math. Soc.* (2), 11 (1912), 411—478 (462, Theorem 37).

so, we take $\alpha = -\sigma' = -\sigma > -1$, $\beta = -\sigma' + \delta = \alpha + \delta > \alpha$. We shall further suppose, as we may, that $\delta < \sigma - \sigma_k$ and $\delta < 1$, so that $\beta - \alpha < 1$.

Now

$$\sum_{n=1}^{\mu-1} \psi_k(n\theta)\, n^{-\sigma}\left(1 - \frac{n}{\mu}\right)^\beta = \sum_{\nu=1}^{\infty} \nu^{-k} \sum_{n=1}^{\mu-1} e(\nu n\theta)\, n^\alpha \left(1 - \frac{n}{\mu}\right)^\beta$$

$$= \sum_{\nu=1}^{\infty} \nu^{-k} S((\nu\theta)) + \sum_{\nu=1}^{\infty} \nu^{-k} \{S(\mu, (\nu\theta)) - S((\nu\theta))\}, \quad \text{.. ......(4·41)}$$

provided that *one* of the series on the right converges. But

$$|S(\mu, (\nu\theta)) - S((\nu\theta))| = |S(\mu, \overline{\nu\theta}) - S(\overline{\nu\theta})| < A\, |\overline{\nu\theta}|^{-(\beta+1)} \mu^{\alpha-\beta},$$

by Lemma N. Also, since $-\beta = \sigma - \delta > \sigma_k$, the series

$$\Sigma\, \nu^{-k}\, |\overline{\nu\theta}|^{-\beta-1}$$

is convergent. Hence

$$\sum_{\nu=1}^{\infty} \nu^{-k}\, |S(\mu, \overline{\nu\theta}) - S(\overline{\nu\theta})|$$

is convergent, and its sum tends to zero (like $\mu^{\alpha-\beta}$) as $\mu \to \infty$. It now follows from (4·41) that

$$\Sigma\, \nu^{-k} S((\nu\theta)) \quad \text{..........................................(4·42)}$$

converges, and that the series

$$\Sigma\, \psi_k(n\theta)\, n^{-\sigma}$$

is summable $(C, \beta)$, *i.e.* $(C, -\sigma' + \delta)$, the sum being given by (4·42). This completes the proof of Theorem 7.

## XXVIII. *Free Paths in a Non-uniform Rarefied Gas with an Application to the Escape of Molecules from Isothermal Atmospheres.*

By J. E. Jones.

1851 Exhibition Senior Research Student, Trinity College, Cambridge.

[*Communicated* by Mr R. H. Fowler.]

[*Read* 5 February 1923.]

§ 1. *Introduction.* Various writers have pointed out that on the basis of the kinetic theory, certain molecules moving in the outer reaches of an atmosphere may, as a result of a series of favourable collisions, receive sufficiently high velocities to take them out of the planet's gravitational field. On this view, the atmosphere of a planet is subject to a continual dissipation. The question of primary interest is the rate at which dissipation occurs, for this decides whether or not a planet may be regarded as retaining its atmosphere.

A simple treatment of the problem has been given by Jeans*, who obtains a formula for the rate of escape of molecules past any atmospheric level. According to this formula, the number of escaping molecules increases when the height of the level increases, as of course it should. But the result depends on the particular atmospheric distribution of molecular density there considered. If, instead, a corrected formula, recently given by Milne†, had been used, Jeans' formula would have indicated the anomalous result that the higher the level considered the less the number of molecules which escape past it, and if the level considered be taken at an infinite distance, as it logically should, the formula would lead to a zero result. This difficulty is avoided by choosing an arbitrary height which is virtually regarded as the ceiling of the atmosphere, so that any molecule passing it with a velocity sufficient to escape from the gravitational field may be regarded as definitely lost. This method involves the neglect of all collisions beyond the arbitrary height and the disregard of all losses from the atmosphere above it. It can hardly be regarded as satisfactory for these considerations require the arbitrary level to be as high as possible, while, as Jeans‡ himself points out, other considerations require it to be as low as possible.

Moreover, the method by which the molecules escaping beyond this arbitrary ceiling are enumerated involves the use of a formula, which is shown in the present paper to be valid only in a gas of uniform density. The value of the molecular density used in the formula is that appropriate to the position of the ceiling, but since, by hypothesis, the ceiling is so high that the atmosphere above it can be neglected, the free paths of the molecules in this region must be so enormous (hundreds, possibly, thousands of kilometres) that there is a sensible change of density along a free path, and a mathematical treatment can only be regarded as satisfactory which takes this into account. This is especially necessary in the case of escaping molecules, since they move from parts of the atmosphere where the density is finite to parts where it is zero.

In order to remove some of the arbitrariness which attaches to the method given by Jeans, Milne§ has recently considered the mechanism of escape in more detail. He has investigated

* Jeans, *Dynamical Theory of Gases*, 3rd edition, 1921, p. 342.

† Milne, *Trans. Camb. Phil. Soc.*, Vol. xxii, p. 483, 1923. Account is taken of the gravitational attraction of the atmosphere itself.

‡ Jeans, *loc. cit.*, p. 344.

§ Milne, *loc. cit.*

the chance of escape of a molecule from the atmosphere without a further collision by introducing the assumption that all the other molecules of the atmosphere are at rest. This leads to the conception of a 'surface level' from below which escape is impossible and to the idea of 'cones of escape' which open continuously from zero angle at the surface level to 90° at infinity.

Although Milne has thus considered the conditions for escape in more detail than hitherto, he has yet adopted the same method of enumerating the numbers which cross a given surface. He is thus led to the result that the number of escaping molecules rises from zero at the surface layer to a maximum and then falls to zero again at infinity. But since the escape of a molecule implies its removal to an infinite distance, the function which expresses the loss from below a given level must increase continuously with increasing height and will strictly give the correct result only in the limit when the level considered approaches infinity.

In the present paper, the problem has been considered from a somewhat different point of view. It has thus been found possible to consider the mechanism of escape without the restriction, introduced by Milne, that all the molecules of the atmosphere are at rest; in fact, no assumption about the molecular velocities has been made beyond their distribution according to the Maxwellian law. The method has involved a discussion of free paths in a non-uniform rarefied gas, where the free path of a molecule is a function not only of its velocity but also of its origin and its direction of motion. General formulae for the free paths have been obtained and, for purposes of illustration, applied to the earth's outer atmosphere. It is shown that, under certain conditions, investigated in detail, a molecule may have an infinite free path, that is, may escape from the atmosphere. For this to occur, the molecule must have its last collision above a certain critical height and must subsequently move within a certain 'cone of escape' appropriate to the point of collision.

The same methods have been used to show that the usual formula for the number of molecules of specified velocities crossing a plane is not applicable when the gas is rarefied and non-uniform, and so cannot legitimately be used to calculate the loss of an atmosphere. The dissipation has been calculated in this paper not by using the formula just referred to, but by considering the collisions in each element of volume of the upper atmosphere and by enumerating those which result in one of the molecules having a velocity of such magnitude and direction as to satisfy all the conditions for escape. The molecules thus lost from this evaporation region we may suppose replaced by diffusion across the critical level from the lower parts of the atmosphere.

According to the methods of this paper, the rate of loss is proportional to the first or the second power of the basic molecular concentration according as the critical level is free or fixed. In general this level sinks as the dissipation proceeds and only becomes fixed when it has descended to the surface of the planet. For this reason the character of the escape of a gas from a planet depends on whether or not it exists alone. Both cases have been considered.

In all cases a gas once having been present never escapes completely. The time taken for the gas to reach such a state of attenuation as to preclude the possibility of detection is in all cases longer than the time given by Jeans for complete escape. Thus the time necessary to reduce the molecular density by $10^5$ is about 10 times as long as that given by him for total escape. All those atmospheric constituents which Jeans considers as being retained by the various planets will therefore be retained according to the present method and so his main conclusions regarding the constituents of the atmospheres of the planets of the solar system are unaffected. There are cases, however, where the retention of a gas is open to doubt according to previous work and the present method of calculation is then necessary. This applies, for instance, to the case of helium on Mars which at certain temperatures would definitely be lost by the more approximate method, but may be regarded as present according to this investigation. This case is considered in some detail and numerical values are given.

The general agreement between these two methods in the case of isothermal atmospheres further justifies us in assuming that the results of Milne obtained for stellar atmospheres of non-uniform temperatures would also be confirmed by the present more detailed analysis.

§ 2. *The Free Paths of Molecules in a Non-uniform Gas.* The only information usually necessary about the free paths of molecules in a gas is the length of the 'mean free path.' While this statistical mean has a definite significance in the case of a uniform gas, where the most probable distance travelled by a molecule between collisions is independent of the direction of its motion, its meaning becomes obscure in the case of a gas such as exists in the upper parts of an atmosphere where there is large variation of molecular density. The path will depend on the direction of motion and may, in fact, be infinite in some directions, while finite in all others. We are therefore led to consider the general problem of the free paths of individual molecules in a gas of non-uniform density. To this end, we confine our attention to a particular molecule moving with known velocity and enquire into its chance of a collision with any other molecule at every point of its path. This has already been done by Tait* in the case of a uniform gas, but the same method can easily be applied to the more general case.

For simplicity, we suppose the molecules of the gas to be rigid elastic spheres. Their diameter we denote by $\sigma$ and their mass by $m$; the number per unit volume we denote by $\nu$.

Consider a molecule of velocity $c$. The chance of a collision with a second molecule of velocity $c'$ in an element of time $\delta t$ is equal to the number of molecules of this kind contained in a cylinder of base area $\pi\sigma^2$ and of length $V\delta t$, where $V$ is the relative velocity. If the direction of $c'$ with respect to that of $c$ be described by the usual Eulerian angular coordinates $\theta$ and $\phi$, then the number of molecules in this cylinder, having velocities between $c'$ and $c'+dc'$ and moving in directions lying between $\theta$ and $\theta+d\theta$, $\phi$ and $\phi+d\phi$, is given by

$$\nu\left(\frac{hm}{\pi}\right)^{\frac{3}{2}} e^{-hmc'^2} c'^2 dc' \sin\theta\, d\theta\, d\phi \times \pi\sigma^2 V\delta t, \quad \text{.......................(2·01)}$$

where $h$ is inversely proportional to the absolute temperature, being given by $2h = 1/kT$.

The velocity $V$ is given, of course, by

$$V^2 = c^2 + c'^2 - 2cc'\cos\theta. \quad \text{..................................(2·02)}$$

If we denote by $\Theta(c)\,\delta t$ the chance of a collision with *any* other molecule in the same interval of time, we have

$$\Theta(c) = \frac{\nu\sigma^2 (hm)^{\frac{3}{2}}}{\pi^{\frac{1}{2}}} \iiint e^{-hmc'^2} c'^2 V dc' \sin\theta\, d\theta\, d\phi, \quad \text{....................(2·03)}$$

the limits of integration being 0 to $\infty$ in the case of $c'$, 0 to $\pi$ in the case of $\theta$ and 0 to $2\pi$ in the case of $\phi$. Tait† found that this expression reduced to

$$\Theta(c) = \frac{\pi^{\frac{1}{2}}\nu\sigma^2}{chm}\psi(c\sqrt{hm}), \quad \text{................................(2·04)}$$

where

$$\psi(x) = xe^{-x^2} + (2x^2+1)\int_0^x e^{-y^2}\, dy. \quad \text{..........................(2·05)}$$

In passing, we may note that the value of $\Theta(c)$ in the case of a stationary molecule is $2\pi^{\frac{1}{2}}\nu\sigma^2/(hm)^{\frac{1}{2}}$, so that if a molecule of hydrogen at normal temperature and pressure were once

* Tait, *Edin. Trans.*, 1886.

† Tait, *loc. cit.*; Jeans, *Dynamical Theory of Gases*, 3rd edition, p. 255.

at rest it would remain so on the average only for $10^{-10}$ seconds. The chance of collision increases continuously with increasing velocity, becoming infinite for an infinite velocity, as is otherwise obvious.

The chance of a collision of a molecule of velocity $c$ with any other molecule in describing any element of its path $\delta s$ is clearly $\Theta(c)\frac{\delta s}{c}$ and hence the length of its path is given by

$$\int_{s_0}^{s} \frac{\pi^{\frac{1}{2}}\sigma^2\psi(c\sqrt{hm})\,\nu ds}{c^2hm} = 1, \qquad (2{\cdot}06)$$

where in a non-uniform gas $\nu$ is a function of $s$. When $s_0$ is known, this is an equation to determine $s$ and therefore the length of the path $\lambda$.

In the absence of an external field of force, the path of the molecule will be rectilinear and its velocity will remain unchanged throughout its path. In this case we have

$$\int_{s_0}^{s} \nu ds = \frac{c^2hm}{\pi^{\frac{1}{2}}\sigma^2\psi(c\sqrt{hm})}, \qquad (2{\cdot}07)$$

which determines a length $\lambda_c$ for each value of the velocity $c$. In the particular case of uniform density the formula simplifies to

$$\lambda_c = \frac{c^2hm}{\pi^{\frac{1}{2}}\nu\sigma^2\psi(c\sqrt{hm})} = \frac{c}{\Theta(c)}. \qquad (2{\cdot}08)$$

Another case of some interest is when $c\sqrt{hm}$ is large. This is the same as $c$ large compared with the mean velocity of thermal agitation $C$, for $hmC^2 = 3/2$. In this case

$$\frac{\psi(c\sqrt{hm})}{c^2hm} \to 2\int_0^{\infty} e^{-y^2}dy$$

$$= \sqrt{\pi},$$

and formula (2·07) becomes

$$\int_{s_0}^{s} \nu ds = \frac{1}{\pi\sigma^2}. \qquad (2{\cdot}09)$$

This equation merely gives the length of that cylinder of base area $\pi\sigma^2$ which contains on the average one molecule. It gives the upper limit to the length of the free path and is evidently the same as that which would be obtained if one molecule was supposed to be moving among a number of other molecules at rest. In the case of uniform density this again reduces to

$$\lambda = \frac{1}{\pi\nu\sigma^2}, \qquad (2{\cdot}10)$$

which is a well-known result.

§ 3. *Probability of a Free Path of given Length.* The usual formula for the probability that a molecule will describe a path of given length also needs generalisation in the case of a non-uniform and rarefied gas such as we consider in the present paper. We have found in the preceding section that there is a certain length $\lambda_c$ associated with a molecule moving with a velocity $c$ in a known direction. It does not follow that every such molecule will actually move through this distance. This length must rather be regarded as the average path traversed by a large number of molecules leaving the same point in the same direction with the same velocity. In general, the lengths of the paths will be distributed about $\lambda_c$ according to some definite law. It is now our purpose to investigate the nature of this law.

Let us denote by $f$ the probability that a molecule moving with velocity $c$ in a direction of angular coordinates $\theta$ and $\phi$ shall describe a free path at least equal to $l$. The chance of a collision in describing a further distance $dl$ beyond the length $l$ is $dl\,.\,\Theta(c)/c$ and hence the probability that the molecule will describe a path of length $l + dl$ is

$$f\,.\left(l - dl\,\frac{\Theta(c)}{c}\right).$$

But by definition this must be the same as $f(l + dl)$. We deduce that

$$\frac{\partial f}{\partial l} = -f\,\frac{\Theta(c)}{c}. \qquad \text{.......(3·01)}$$

Now $\Theta(c)$ involves $\nu$ and therefore the right-hand side of the above equation is a function of $l$, of the angular coordinates $\theta$, $\phi$ of the molecule's motion, as well as of the coordinates of the starting point of the free path. Writing

$$\nu = g(x_0, y_0, z_0, l, \theta, \phi),$$

where $x_0$, $y_0$, $z_0$ are the coordinates of the starting point, and putting

$$\frac{\Theta(c)}{\nu c} = \theta(c) = \frac{\pi^{\frac{1}{2}}\sigma^2\,\psi(c\sqrt{hm})}{c^2 hm}, \qquad \text{.......(3·011)}$$

we have

$$f = e^{-\theta(c)\int g\,dl}, \qquad \text{.......(3·02)}$$

on using the condition that $f(0) = l$. It follows that the probability of a path of length between $l$ and $l + dl$ is

$$-\,\theta(c)\,g\,dl\,e^{-\theta(c)\int g\,dl}. \qquad \text{.......(3·03)}$$

When $\nu$ is constant, equation (3·02) reduces to

$$f = e^{-\nu\theta(c)\,l} = e^{-\frac{\Theta(c)\,l}{c}} = e^{-l/\lambda_c}, \qquad \text{.......(3·04)}$$

while formula (3·03) becomes

$$-\frac{dl}{\lambda_c}\,e^{-l/\lambda_c}. \qquad \text{.......(3·05)}$$

These are the formulae given by Jeans*. That $\lambda_c$ is the average path is easily verified, for

$$\int_0^\infty l e^{-l/\lambda_c}\,dc = \lambda_c\,.$$

§ 4. *Free Paths in an Upper Atmosphere.* When account is taken of the variation of a planet's gravitational attraction with increasing height, it is found that the molecular density of any constituent of its atmosphere at a distance $r$ from its centre is given by†

$$\nu = \nu_0\,e^{-2hmg\frac{a(r-a)}{r}}, \qquad \text{.......(4·01)}$$

where $\nu_0$ is its value at the base of the isothermal part of the atmosphere‡, $h$ is inversely proportional to the temperature ($2h = 1/kT$), $g$ is the value of gravity at the surface and $a$ the radius of the planet. This formula, however, gives a finite density at an infinite distance and under such conditions no molecule ever could be said to have escaped from the atmosphere. If account be taken of the gravitational attraction of the atmosphere itself on its outer fringes, it is found that at very large distances the molecular density falls off according to an inverse square law. Milne§ suggests as an appropriate formula applicable at all distances, the law

$$\nu = \nu_0\left(\frac{r_0}{r}\right)^2 e^{-q_0'\left(1-\frac{r_0}{r}\right)}, \qquad \text{.......(4·02)}$$

* Jeans, *loc. cit.*, p. 347.

† Jeans, *loc. cit.*, p. 343.

‡ This is not quite the same as the base of the stratosphere. See Chapman and Milne, *Quart. Journal Roy. Met. Soc.*, Vol. 46, 1920.

§ Milne, *loc. cit.*, p. 491, equation 25.

where $r_0$ refers to any convenient level in the atmosphere and

$$q_0' = 2hmg\, r_0 - 2 = q_0 - 2. \quad \text{.................................(4·03)}$$

This formula we propose to adopt in the present investigation. We shall take $r_0$ to be the radial distance of the base of the isothermal part of the atmosphere.

Suppose now that a molecule, having collided at a point distant $r$ from the centre, moves with a velocity $c$ in a rectilinear path making an angle $\theta$ with the radial direction at that point. Strictly its subsequent path will be hyperbolic if, as in the cases we consider, the magnitude of $c$ is such that the kinetic energy is greater than the gravitational potential. Since such velocities are large, the curvature of the path will at any rate be small and we may regard the path as rectilinear.

If $s$ denote the path described from the origin of its new velocity and $R$ denote the radial distance corresponding to $s$, we have

$$R = (r^2 + 2rs \cos\theta + s^2)^{\frac{1}{2}},$$

and the length of the free path is given by

$$\int_0^s \nu(R)\, ds = \frac{c^2 hm}{\pi^{\frac{1}{2}} \sigma^2 \psi(c\sqrt{hm})}. \quad \text{.................................(4·04)}$$

This equation is true only when $c$ remains constant along the whole path, but we may apply it without sensible error to the case of molecules moving in an upper atmosphere. When the path is sufficiently long to make an appreciable change in $c$, the molecule has by that time reached parts of the atmosphere where the molecular density is negligible.

A further simplification can be introduced by writing

$$R = r + s\cos\theta. \quad \text{.......................................(4·05)}$$

This is equivalent to neglecting the curvature of the layers of equal density through which a molecule may be supposed to pass. The equation (4·04) to determine the free path then becomes

$$\int_0^s \frac{e^{q_0' \frac{r_0}{r + s\cos\theta}}}{(r + s\cos\theta)^2}\, ds = \frac{e^{q_0'} c^2 hm}{\pi^{\frac{1}{2}} \nu_0 \sigma^2 r_0^2 \psi(c\sqrt{hm})},$$

which reduces to

$$e^{q_0' \frac{r_0}{r}} - e^{q_0' \frac{r_0}{R}} = \frac{q_0' e^{q_0'} \cos\theta\, c^2 hm}{\pi^{\frac{1}{2}} \nu_0 \sigma^2 r_0 \psi(c\sqrt{hm})}. \quad \text{..........................(4·06)}$$

When $r$, $c$ and $\theta$ are given, this equation determines $R$ and therefore $s$, the free path, from the relation

$$s = (R - r)\cos\theta.$$

Or again, if $r$ and $\theta$ are given, it can be regarded as an equation to give the velocity necessary to describe a path of given length. In particular, the relation which must exist between the three quantities $r$, $c$ and $\theta$ in order that a molecule may escape from the atmosphere is

$$e^{q_0' \frac{r_0}{r}} - 1 = \frac{q_0' e^{q_0'} \cos\theta\, c^2 hm}{\pi^{\frac{1}{2}} \nu_0 \sigma^2 r_0 \psi(c\sqrt{hm})}. \quad \text{..........................(4·07)}$$

Moreover, if in this equation we put $\theta = 0$ and $c = \infty$, we have an equation to determine the lowest level from which escape is possible. Thus, using $r_c$ to denote the height of that critical layer, we have

$$e^{q_0' \frac{r_0}{r_c}} - 1 = \frac{q_0' e^{q_0'}}{\pi \nu_0 \sigma^2 r_0}. \quad \text{.................................(4·08)}$$

We note that the molecular concentration at the critical level is given by

$$\nu_c = \nu_0 \left(\frac{r_0}{r_c}\right)^2 e^{-q_0'\left(1-\frac{r_0}{r_c}\right)} = \left(\frac{r_0}{r_c}\right)^2 \left\{\nu_0 e^{-q_0'} + \frac{q_0'}{\pi\sigma^2 r_0}\right\},$$

and so, when $q_0'$ is large, by $q_0'/\pi\sigma^2 r_0$ approximately. The order of this expression is determined almost entirely by that of $\sigma^2$ and $r_0$, and as these are of the same order for all molecules, we find the interesting result that on any planet the molecular density at the critical layer is always of the same order, whatever its value at the base or whatever the constitution of the atmosphere. The order is $10^8$ in the case of the earth.

Corresponding to any value of $r$ greater than $r_c$, there is a range of values of $\theta$ for which escape is possible. The limiting value will clearly be a function of $r$, and so we denote it by $\theta_r$. It is given by

$$\cos\theta_r = \frac{\pi\nu_0\sigma^2 r_0}{q_0' e^{q_0'}}\left(e^{q_0'\frac{r_0}{r}} - 1\right), \quad \ldots\ldots\ldots\ldots\ldots\ldots(4{\cdot}09)$$

or, using equation (4·08), by

$$\cos\theta_r = \frac{e^{q_0'\frac{r_0}{r}} - 1}{e^{q_0'\frac{r_0}{r_c}} - 1}. \quad \ldots\ldots\ldots\ldots\ldots\ldots(4{\cdot}10)$$

For escape from this height a molecule must move within a cone of semi-angle $\theta_r$. This cone we shall refer to as the *cone of escape*.

For purposes of illustration, we now apply these formulae to the case of the earth's outer atmosphere; here $r_0 = 6{\cdot}39 \times 10^8$ cms., the temperature is usually considered* to be $-54^\circ$ C., and assuming the outer constituent to be hydrogen† for which‡ $\frac{k}{m} = 4{\cdot}127 \times 10^7$, we have

$$q_0' = 2hmgr_0 - 2 = \frac{mgr_0}{kT} - 2$$
$$= 67{\cdot}37.$$

We may, therefore, neglect the unit terms in equations (4·08) and (4·10), and write

$$\cos\theta_r = e^{q_0'\left(\frac{r_0}{r} - \frac{r_0}{r_c}\right)}. \quad \ldots\ldots\ldots\ldots\ldots\ldots(4{\cdot}11)$$

Again, for hydrogen we have§ $\sigma = 2{\cdot}72 \times 10^{-8}$ cms., while $\nu_0 = 1{\cdot}89 \times 10^{19}$, and we thus find $r_c = 1{\cdot}238 r_0$. The critical layer is thus at a height of 1521 kilometres above the base of the stratosphere. The 'cone of escape' opens very rapidly above this height, as the values in the accompanying table testify.

TABLE I.

*Variation of Cone of Escape with Height (Earth's Hydrogen).*

| $\theta_r$ | $r/r_0$ | $r - r_0$ (in kilometres) |
|---|---|---|
| 0 | 1·238 | 1521 |
| 25° | 1·240 | 1533 |
| 45° | 1·246 | 1571 |
| 65° | 1·254 | 1623 |
| 85° | 1·296 | 1891 |

* Jeans, *loc. cit.*, p. 345.

† There is however some doubt as to whether this is so; see Chapman and Milne, *Quart. Journ. Roy. Met. Soc.*, Vol. 46, 1920.

‡ Jeans, *loc. cit.*, p. 119.

§ Jeans, *loc. cit.*, chap. XIV.

Chapman and Milne, *loc. cit.* (base of stratosphere 20 kms. high).

The cone of escape thus opens from 0° to 45° in a distance of 50 kilometres and from 0° to 85° in a distance of 370 kilometres.

It is not to be implied that all molecules moving in directions within the 'cone of escape' can eventually escape from the atmosphere without collision. We have already seen that the velocity must be sufficiently large to take it out of the earth's gravitational field. But this velocity may not be sufficiently high to ensure that its subsequent motion will be free from collision. For this condition to be satisfied, a certain velocity $\mathscr{C}_{r,\theta}$ (a function of $r$ and $\theta$) must be exceeded. This value of $\mathscr{C}_{r,\theta}$ is easily obtained from equation (4·07) and is given by

$$\frac{\pi^{\frac{1}{2}}\mathscr{C}^2_{r,\theta}hm}{\psi(\mathscr{C}_{r,\theta}\sqrt{hm})}=\frac{\left(e^{q_0'\frac{r_0}{r}}-1\right)\pi\nu_0\sigma^2 r_0}{q_0' e^{q_0'}\cos\theta}$$

$$=\frac{\cos\theta_r}{\cos\theta}\,. \qquad\qquad (4{\cdot}12)$$

The value of $\psi(x)$ can only be obtained for particular values of $x$ by quadrature, but a sufficient number of values have been given by Tait* and reproduced by Jeans† to deduce the value of $\mathscr{C}_{r,\theta}\sqrt{hm}$ for particular values of the right-hand side. A few values of $\psi(x)$ and $x^2/\psi(x)$ are given in the accompanying table.

TABLE II.

| $x$ | $x^2$ | $\psi(x)$ | $x^2/\psi(x)$ |
|---|---|---|---|
| ·1 | ·01 | ·20066 | ·0498 |
| ·5 | ·25 | 1·08132 | ·2312 |
| 1·0 | 1·0 | 2·60835 | ·3835 |
| 1·5 | 2·25 | 4·86713 | ·4624 |
| 2·0 | 4·00 | 7·97536 | ·5016 |
| 2·5 | 6·25 | 11·96402 | ·5221 |
| 3·0 | 9·00 | 16·83830 | ·5344 |
| 5·0 | 25·00 | 45·186 | ·5534 |
| 10·0 | 100·00 | 178·086 | ·5620 |
| $\infty$ | $\infty$ | $\infty$ | ·5642 $(=\pi^{-\frac{1}{2}})$ |

Thus at the level for which $\theta_r=45°$, the minimum value of $\mathscr{C}_{r,\theta}\sqrt{hm}$ (being that appropriate to $\theta=0$) is 1·09, which at a temperature of −54° C. corresponds to

$$\mathscr{C}_{r,\theta}=1{\cdot}457\times10^5 \text{ cms. sec.}, \quad (\sqrt{hm}=7{\cdot}463\times10^{-6}).$$

For larger velocities than this, there is a corresponding range of directions for which escape is possible, the upper limit to the range being as we have indicated 45°, and this only in the case of infinite velocities. A curve showing the relation between any velocity and the extreme value of $\theta$ for which escape is possible is shown in Fig. 1.

The conditions for escape which we have so far examined can be expressed simply by saying that the end point of a vector ($OP$) drawn from $O$ to represent the velocity must lie within a certain surface of revolution. This surface is obtained by revolving the curve $AV$ of Fig. 1 about the vertical.

We must be quite clear, however, that we have so far only been concerned with the possibility of escape *without a collision*. In general, there is another factor affecting escape and that is the

* Tait, *loc. cit.*

† Jeans, *loc. cit.*, Appendix B.

gravitational potential of the planet. A certain minimum velocity $c_r$ must be exceeded to ensure that the molecule will escape from the gravitational field. It may be that this velocity is less than the minimum value of $\mathscr{C}_{r,\theta}$ or it may be that it is more. In the latter case, the region in which the end point $P$ of the vector must lie is no longer $VAV'$ but only that portion of it above the spherical cap of radius $c_r$.

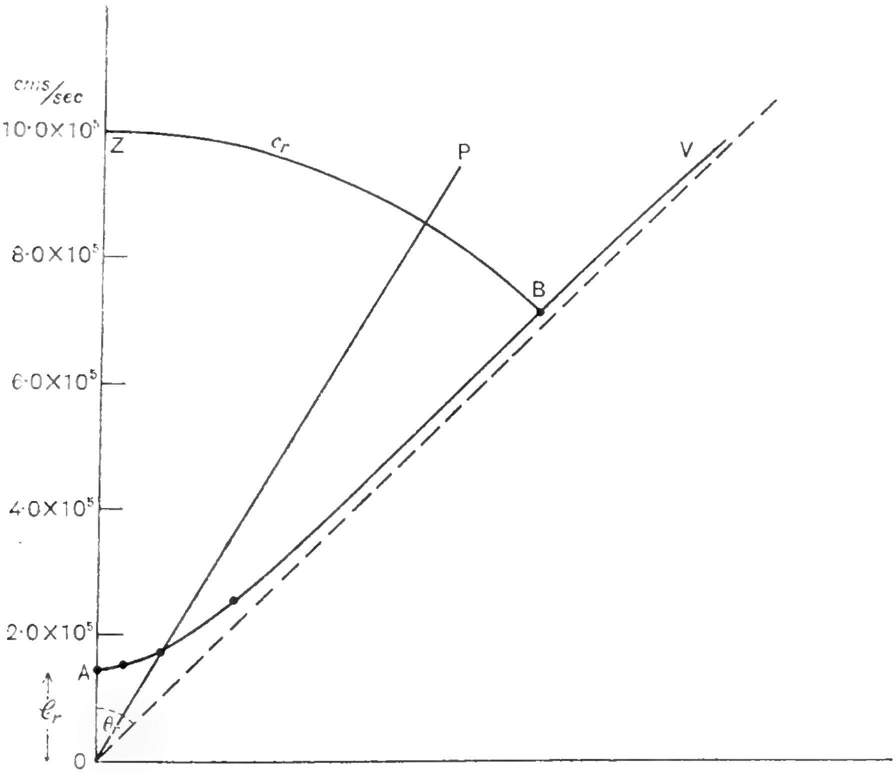



Fig. 1. Velocities necessary for escape in different directions ($\theta_r = 45°$).

In the actual case considered above, the gravitational attraction is very large and the corresponding value of $c_r$ is $1{\cdot}001 \times 10^6$ cms./sec. This velocity is larger than $\mathscr{C}_{r,\theta}$ through a range of $\theta$ of 44·5° and is therefore the dominating factor. We shall return to a discussion of the relative magnitudes of $c_r$ and $\mathscr{C}_{r,\theta}$ in the general case later in the paper.

If after a collision in the upper atmosphere a molecule moves in a direction outside the 'cone of escape,' its free path will be of finite length, however great its velocity. The length of this path will of course vary according to the direction, being a minimum for a molecule moving vertically downwards. For purposes of illustration, the free paths of molecules leaving a given point in the upper atmosphere with various velocities in different directions have been worked out and are represented in the accompanying figure. The particular point of departure considered is that for which the 'cone of escape' has an angle of 45° ($r = 1{\cdot}246\, r_0$). The formula used is

$$\frac{e^{q_0' \frac{r_0}{r}} - e^{q_0' \frac{r}{R}}}{e^{q_0' \frac{r_0}{r}} - 1} = \frac{\pi^{\frac{1}{2}} c^2 hm}{\psi(c\sqrt{hm})} \frac{\cos\theta}{\cos\theta_r}, \quad \text{.............................} (4{\cdot}13)$$

derived from equations (4·06) and (4·09); except for values of $r$ which are large compared with $r_0$, this equation can be simplified to

$$q_0'\left(\frac{r_0}{R}-\frac{r_0}{r}\right)=\log_e\left(1-\frac{\pi^{\frac{1}{2}}c^2hm}{\psi(c\sqrt{hm})}\frac{\cos\theta}{\cos\theta_r}\right). \qquad (4{\cdot}14)$$

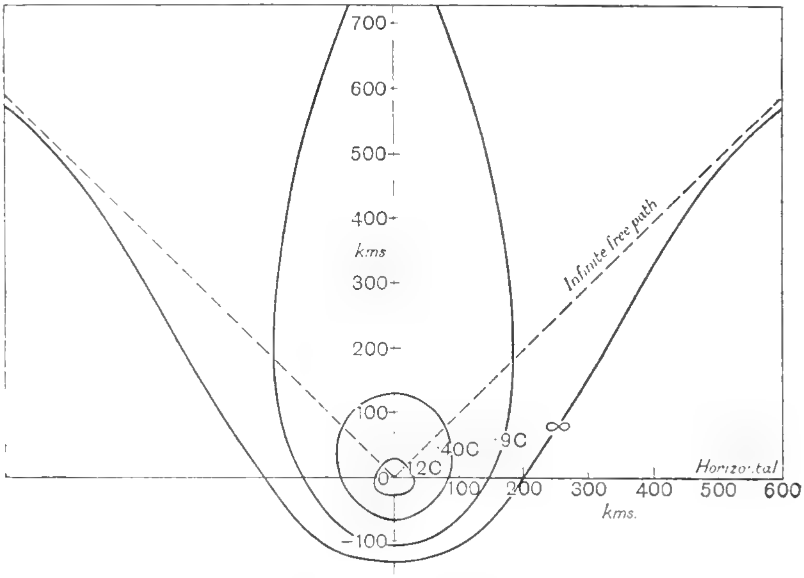



Fig. 2. Free paths of molecules leaving a given point in the upper atmosphere with various velocities in different directions.
($C$ = mean velocity of thermal agitation = $1{\cdot}640\times10^5$ cms./sec. at $-51^\circ$ C.)

§ 5. *The Number of Molecules of specified Velocities which cross a given Plane.* The main object of the present paper is to evaluate the rate of loss of molecules from the earth's atmosphere, and as the calculations of previous writers have all depended on the use of a formula for the number of molecules of given velocities which cross a given plane, it seems worth while to enquire whether this formula, like the formula for the length of the free path, requires correction in a gas which is both non-uniform and rarefied. The formula in question is*

$$\nu\left(\frac{hm}{\pi}\right)^{\frac{3}{2}}e^{-hm(u^2+v^2+w^2)}\,c\cos\theta\,du\,dv\,dw\,dS. \qquad (5{\cdot}01)$$

This purports to give the number of those molecules whose component velocities lie between $(u, v, w)$ and $(u+du, v+dv, w+dw)$ which cross an element of area $dS$ per unit time. It is obtained by counting the number of such molecules which at any time lie in a cylinder of base $dS$ and length $c\,dt$ drawn in the direction of the resultant velocity $c$, $\theta$ being the angle between the axis of the cylinder and the normal to $dS$.

It is to be observed that the value of $\nu$ in (5·01) is that appropriate to the position of $dS$. In a rarefied gas, however, molecules will have travelled large distances before reaching the plane, and if the gas is of variable density, it follows that such molecules have arrived from parts of the gas where the molecular density $\nu$ is appreciably different from that at the position of the plane. In order to investigate whether this will introduce any modification in the expression (5·01), we propose to develop a more fundamental method of calculation than that previously used.

* Jeans, *loc. cit.*, p. 343; Milne, *loc. cit.*, equation (28).

We consider only the case of a gas which has reached a steady state, that is one in which the molecular density and the distribution of velocities about their mean do not vary with the time. We make no assumption about the nature of the velocity distribution function. It need not be Maxwellian, although in the applications which we are likely to make of the results this will usually be the case.

All the molecules which cross an element of area $dS$ with velocities lying between $(u, v, w)$ and $(u+du, v+dv, w+dw)$ must have had their origin, that is have had their last collision, in a cylinder of infinite length drawn in the direction of the velocity $c$ on the base $dS$. We consider those which collided last in a thin slab of this cylinder of thickness $dx$. What we require first is the number of collisions in this element of volume such that one of the molecules *after* collision has a velocity lying in the specified range. Since the gas is in a steady state, this number is equal to the number of collisions in which one of the molecules had the specified velocity *before* collision. If $f(u, v, w)$ be the fraction of the molecules in any element of volume possessing component velocities between $(u, v, w)$ and $(u+du, v+dv, w+dw)$, the number of collisions of the type considered per unit time is

$$\nu f(u, v, w)\, du\, dv\, dw\, \Theta(c)\, dx\, dS, \qquad (5\cdot02)$$

since $\Theta(c)$ is the chance of collision per unit time of any molecule moving with velocity $c$. In the notation of § 3 we may rewrite this expression in the form

$$\nu^2 f(u, v, w)\, c\, \theta(c)\, du\, dv\, dw\, dx\, dS. \qquad (5\cdot03)$$

The fraction of these which reach the area $dS$ will depend on the distance of the slab from the area. Let us suppose that the distance measured along the axis is equal to $l$. If $\theta$ is the angle between the axis and the normal to $dS$, we have then $dx = dl\cos\theta$. In the section just referred to, § 3 above, we found that the probability that a molecule moving with velocity $c$ should travel along a path at least equal to $l$ was given by $e^{-\theta(c)\int\nu\, dl}$. Hence the number of molecules which leave the slab $dx\, dS$ per unit time with the prescribed velocity and reach the area $dS$ is

$$\nu^2 f(u, v, w)\, c\cos\theta\, \theta(c)\, e^{-\theta(c)\int\nu\, dl}\, du\, dv\, dw\, dl\, dS, \qquad (5\cdot04)$$

and hence the total number arriving at the area $dS$ from all parts of the cylinder is obtained by integrating the expression with respect to $l$ from 0 to $\infty$. We have

$$dN = f(u, v, w)\, c\cos\theta\, \theta(c)\, du\, dv\, dw\, dS \int_0^\infty \nu^2 e^{-\theta(c)\int\nu\, dl}\, dl. \qquad (5\cdot05)$$

In any given problem, $\nu$ will be a known function of $l$.

In the case of uniform density, the formula reduces to

$$dN = \nu f(u, v, w)\, c\cos\theta\, du\, dv\, dw\, dS, \qquad (5\cdot06)$$

the usual formula; but it is only in this special case that the two are identical. It may, however, be used with sufficient accuracy in any gas of normal density whether uniform or not. For in this case we have

$$\int_0^\infty \nu^2\theta(c)\, e^{-\theta(c)\int\nu\, dl}\, dl = \int_0^\infty \nu e^{-y}\, dy, \qquad (5\cdot07)$$

where

$$y = \theta(c)\int_0^l \nu\, dl.$$

Now if, as we have supposed, $\nu$ is large, $y$ becomes very large even for values of $l$ which are very minute on ordinary standards of measurement, so that any variation in $\nu$ in finite distances is rendered negligible by the factor $e^{-y}$. It follows that in such cases the integral $\int_0^\infty \nu e^{-y}\, dy$ may be evaluated as though $\nu$ were constant, leading of course to the value $\nu$. In the case of a gas like the upper atmosphere, the formula (5·06) is not correct and must be replaced by (5·05).

A simplification can be introduced into formula (5·05) by supposing that all molecules which leave any point with a given velocity $c$ travel the same distance $\lambda_c$, given by equation (2·07). All the molecules which cross $dS$ with the prescribed velocity will then have had their last collision in a cylinder of base $dS$ and length $\lambda_c$ drawn in the direction of $c$. We infer that the number required will then be obtained by integrating the expression (5·03) throughout the volume of the cylinder. We find

$$dN = f(u, v, w)\, c \cos\theta\, \theta(c)\, du\, dv\, dw\, dS \int_0^{\lambda_c} \nu^2\, dl. \quad \text{.................(5·08)}$$

This formula, like (5·05), reduces to the usual expression (5·06) whenever the molecular density is uniform or does not change appreciably along a free path.

To illustrate the difference between the usual formula (5·06) and the formula (5·05) we consider a particular case. Let us conceive of a hypothetical atmosphere in which the density falls off linearly from a finite known value at its base to zero at its ceiling, so that

$$\nu = \frac{\nu_0 (a - x)}{a}$$

with an obvious notation, and let us calculate the number of molecules of specified velocities which arrive at the ceiling. According to formula (5·06) the number is zero for the molecular density at the ceiling is zero and it is that value which is to be substituted in the formula. On the other hand, formula (5·05) gives, on putting $\cos\theta\, dl = dx$,

$$dN = f(u, v, w)\, c\theta(c)\, du\, dv\, dw\, dS \int_0^a \nu^2 e^{-\frac{\theta(c)\,\nu\,(a-x)^2}{2a\cos\theta}}\, dx$$

$$= f(u, v, w)\, c\theta(c)\, du\, dv\, dw\, dS\, \frac{\nu_0^2}{a^2} \int_0^a y^2 e^{-\frac{\theta(c)\,\nu_0\, y^2}{2a\cos\theta}}\, dy,$$

and so if $a$ is large, as we may suppose it to be, we have

$$dN = f(u, v, w) = \sqrt{\frac{\pi \nu_0 \cos^3\theta}{a\theta(c)}}\, du\, dv\, dw\, dS,$$

where $c$ is the magnitude of the velocity and $\theta$ is the angle between its direction and that of the axis of $x$. The total number of molecules which cross unit area per second is therefore given by

$$N = \iiint f(u, v, w)\, c\, \pi^{\frac{1}{4}} \sqrt{\frac{\nu_0 hm\, c^2 \cos^3\theta}{a\sigma^2 \psi(c\sqrt{hm})}}\, c^2 \sin\theta\, d\theta\, d\phi\, dc,$$

and assuming a Maxwellian distribution function, this becomes

$$N = \frac{4\pi^{\frac{5}{4}}}{5} \left(\frac{\nu_0}{a\sigma^2}\right)^{\frac{1}{2}} \left(\frac{hm}{\pi}\right)^{\frac{3}{2}} \int_0^\infty e^{-hmc^2} c^3 \sqrt{\frac{hm c^2}{\psi(c\sqrt{hm})}}\, dc$$

$$\sim \left(\frac{\nu_0}{a\sigma^2}\right)^{\frac{1}{2}} \frac{1}{(hm)^{\frac{1}{2}}}.$$

If this law of density represented the case of the earth's density, where for hydrogen $\nu_0$ is of the order of $10^{13}$, $\sigma^2$ is of order $10^{-15}$, and $hm$ is of order $10^{-10}$, we should get $N$ to be of order $10^{19} a^{-\frac{1}{2}}$. If $a$ be of the order of the earth's radius $10^8$ cms., we then find that $10^{15}$ molecules cross unit area of the ceiling per second. This is in contrast to the zero result obtained by the ordinary method.

§ 6. *Conditions for the Escape of a Molecule.* Before proceeding to the calculation of the rate of escape of molecules from an atmosphere, it will be convenient to summarise the conditions necessary for escape.

(i) The point of the last collision of the molecule must be higher than a certain critical height, the precise value of which depends on the nature of the gas ($r > r_c$).

(ii) The direction of motion after collision must fall within a certain 'cone of escape,' the size of which depends on the position of the point of collision ($\theta < \theta_r$).

(iii) The velocity must be sufficiently large to take the molecule out of the earth's gravitational field $\left(c > c_r, \text{ where } c_r^2 > \frac{2ga^2}{r}\right)$.

(iv) The velocity must be sufficiently high to avoid further collisions ($c > \mathscr{C}_{r,\theta}$, where $\mathscr{C}_{r,\theta}$ depends on the point of collision and on the direction of motion).

In general then there will be a stratum in which $\mathscr{C}_{r,\theta}$ is greater than $c_r$ and in which the sole criterion for escape is the possibility of an infinite path without a collision. The upper boundary of the stratum will be determined by $\mathscr{C}_{r,0} = c_r$ and above it the dominating condition for escape is that the velocity of a molecule shall exceed $c_r$.

§ 7. *The Loss of Molecules from a Simple Isothermal Atmosphere.* The methods of § 4 can now be applied to find how many collisions take place per unit time in an elementary slab of the atmosphere such that after collision one of the molecules satisfies all the conditions mentioned in the preceding section. For the present we shall confine our attention to an atmosphere which consists of one constituent only. Such an atmosphere we shall refer to as a *simple* atmosphere. The work can easily be extended to a mixed atmosphere, as we afterwards show, but to do so at the present stage would introduce unnecessary complications. From equation (5·03) we infer that the number of molecules which in unit time undergo collisions in a spherical shell of thickness $dr$ and thereby receive a velocity, whose magnitude lies between $c$ and $c + dc$ and whose direction makes an angle of $\theta$ to $\theta + d\theta$ with the vertical, is

$$\nu^2 f(u, v, w)\, 2\pi c^3 \theta(c) \sin\theta\, d\theta\, dc\; 4\pi r^2 dr. \quad \text{.......................(7·01)}$$

If the magnitude of $c$ exceeds the values $\sqrt{2ga^2/r}$, ($c_r$) and $\mathscr{C}_{r,\theta}$, and $\theta$ lies within the cone of escape appropriate to the value of $r$, then all the molecules given by (6·01) will escape from the atmosphere. The total number of escaping molecules is therefore given by

$$dL = \nu^2 8\pi^2 r^2 dr \int_{C_r}^{\infty} \int_0^{\theta_r} f(u, v, w)\, c^3 \theta(c) \sin\theta\, d\theta\, dc, \quad \text{..................(7·02)}$$

where we have written $C_r$ for the greater of the two quantities $c_r$ and $\mathscr{C}_{r,\theta}$. In an isothermal atmosphere the distribution may be regarded as Maxwellian, and we put

$$f = \left(\frac{hm}{\pi}\right)^{\frac{3}{2}} e^{-hmc^2}.$$

Introducing this value for $f$ in equation (6·02) and substituting for $\theta(c)$ its value from equation (3·011),

$$\theta(c) = \frac{\pi^{\frac{1}{2}} \sigma^2 \psi(c\sqrt{hm})}{c^2 hm},$$

we find
$$dL = 8\pi\sigma^2 (hm)^{\frac{1}{2}} \nu^2 r^2 dr \int_{C_r}^{\infty} \int_0^{\theta_r} e^{-hmc^2} c\, \psi(c\sqrt{hm}) \sin\theta\, d\theta\, dc$$

$$= 8\pi\sigma^2 (hm)^{\frac{1}{2}} \nu^2 r^2 dr \int_{C_r}^{\infty} e^{-hmc^2} c\, \psi(c\sqrt{hm}) (1 - \cos\theta)\, dc, \quad \text{..............(7·03)}$$

where, as before,
$$\psi(x) = xe^{-x^2} + (2x^2 + 1) \int_0^x e^{-y^2} dy.$$

There is, as explained in § 4, a limiting value of $\theta$ for each value of $c$, so that $\cos\theta$ is a function of $c$. The required relation is contained in equation (4·07), viz.

$$e^{q_0 \frac{r_0}{r}} - 1 = \frac{q_0 e^{q_0} \cos\theta\, c^2 hm}{\pi^{\frac{1}{2}} \nu_0 \sigma^2 r_0 \psi(c\sqrt{hm})}, \qquad \text{(7·04)}$$

but this can be expressed more neatly by using the fact that the lowest value of $c$ for escape from a given level is given by

$$\frac{\psi(\mathscr{C}_r\sqrt{hm})}{\mathscr{C}_r^2 hm} = \frac{q_0 e^{q_0}}{\pi^{\frac{1}{2}} \nu_0 \sigma^2 r_0 \left(e^{q_0 \frac{r_0}{r}} - 1\right)}, \qquad \text{(7·05)}$$

so that for any direction other than the vertical

$$\cos\theta = \frac{\mathscr{C}_r^2}{\psi(\mathscr{C}_r\sqrt{hm})}\,\frac{\psi(c\sqrt{hm})}{c^2}. \qquad \text{(7·06)}$$

We accordingly have

$$dL = 8\pi\sigma^2 (hm)^{\frac{1}{2}} \nu^2 r^2 dr \int_{\mathscr{C}_r\sqrt{hm}}^{\infty} e^{-x^2} x\,\psi(x)\left(1 - \frac{\mathscr{C}_r^2 hm}{\psi(\mathscr{C}_r\sqrt{hm})}\,\frac{\psi(x)}{x^2}\right)dx. \qquad \text{(7·07)}$$

Owing to the nature of $\psi(x)$, this integral is too complicated to be evaluated in the general case except by quadrature, but in certain cases the work can be very much simplified. In the first place we observe that, even for comparatively small values of $x$, the term $\int_0^x e^{-y^2}dy$ occurring in $\psi(x)$ differs very little from its limiting value $\sqrt{\pi}/2$. Thus, when $x = 2$, we find that its value is $(\cdot 99532)\sqrt{\pi}/2$. Furthermore, for such values of $x$, the term $xe^{-x^2}$ is quite small. For the value $x = 2$, we find a value of ·03664. In all cases therefore in which $x$ is larger than 2, very little error will be introduced by supposing that

$$\psi(x) = \frac{(2x^2+1)\sqrt{\pi}}{2}. \qquad \text{(7·08)}$$

We find further support for this approximation in that

$$\int_0^\infty e^{-x^2} x\,\psi(x) = \int_0^\infty e^{-x^2} x^2 dx + \int_0^\infty e^{-x^2}(2x^3 + x)\,dx \int_0^x e^{-y^2}dy$$

$$= \sqrt{\frac{\pi}{2}} = \cdot 7071\sqrt{\pi}.$$

whereas
$$\int_0^\infty e^{-x^2} x(2x^2+1)\frac{\sqrt{\pi}}{2}\,dx = \cdot 75\sqrt{\pi},$$

so that even when integrated through the infinite range, the difference between them is small*.

* It will be as well at this stage to enquire what values of the lower limit $C_r$ we are likely to meet in the application of this work to actual planetary atmospheres. At the critical level, of course, its value is infinite by definition. At higher heights, its value decreases until $C_r$ is equal to the 'gravitational' velocity $c_r$. We then have

$$C_r\sqrt{hm} = \sqrt{2hmgr_0^2/r} = \sqrt{q_0 r_0/r}.$$

Its value therefore depends on the temperature of the atmosphere, on the mass of the molecules, on the gravitational potential of the planet at its surface, and on the level considered ($r$). If we take an extreme case, viz. that of the moon, where the gravitational potential is least, and consider hydrogen molecules for which $m$ is least, we find at a temperature of $-100°$ C.,

$$C_r\sqrt{hm} = 3\cdot 93\sqrt{r_0/r}.$$

Hence even up to a height of $4r_0$ ($r_0$ being the radius of the moon), were it possible to conceive of an atmosphere of such extent, our work would still apply. We may, therefore, safely regard the analysis as applicable to the escape of molecules (as distinct from electrons) from any known planet.

Making this substitution for $\psi(x)$ in equation (7·07) we then find

$$dL = \frac{4\pi^{\frac{3}{2}}\sigma^2\nu^2 r^2 dr}{(hm)^{\frac{1}{2}}}$$

$$\times\left[e^{-hmC_r^2}\left\{\left(\frac{3}{2} - \frac{2\sqrt{\pi}\,\mathscr{C}_r^2 hm}{\psi(\mathscr{C}_r\sqrt{hm})}\right) + hmC_r^2\sqrt{\pi}\left(1 - \frac{\mathscr{C}_r^2 hm}{\psi(\mathscr{C}_r\sqrt{hm})}\right)\right\} + \frac{\mathscr{C}_r^2 hm}{\psi(\mathscr{C}_r\sqrt{hm})}\frac{\sqrt{\pi}}{4}\int_{hmC_r^2}^{\infty}\frac{e^{-y}}{y}dy\right].$$

Now for large values of $x$ we can write*

$$\int_x^{\infty}\frac{e^{-y}dy}{y} = e^{-x}\left\{\frac{1}{x} - \frac{1}{x^2} + \frac{2!}{x^3} - \dots\right\},$$

so that taking the first term only where $x = hmC_r^2$, we have

$$dL = \frac{4\pi^{\frac{3}{2}}\sigma^2\nu^2 r^2 dr}{(hm)^{\frac{1}{2}}}e^{-hmC_r^2}$$

$$\times\left\{\left(\frac{3}{2} - \frac{2\sqrt{\pi}\,\mathscr{C}_r^2 hm}{\psi(\mathscr{C}_r\sqrt{hm})}\right) + hmC_r^2\sqrt{\pi}\left(1 - \frac{\mathscr{C}_r^2 hm}{\psi(\mathscr{C}_r\sqrt{hm})}\right) + \frac{\mathscr{C}_r^2 hm}{\psi(\mathscr{C}_r\sqrt{hm})}\frac{\sqrt{\pi}}{4hmC_r^2}\right\}. \quad \text{.........(7·09)}$$

We note in passing that if in this result $dr$ be replaced by $\frac{1}{\sqrt{2}\,\pi\sigma^2\nu}$, we get a formula for $L$ which is almost identical with that used by Jeans†. What differences there are, are due principally to the introduction here of the idea of 'angles of escape.' Now this operation is obviously equivalent to integrating (7·09) through a length $\frac{1}{\sqrt{2}\,\pi\sigma^2\nu}$ as though the function were constant, and since this length is equal to the mean free path in a gas of uniform density $\nu$, we are led to an interpretation of Jeans' formula. Instead of giving the escape of molecules past a given level from *all* parts of the atmosphere below it, this gives only the escape from a certain slab of the atmosphere of a certain definite thickness; this thickness is equal to the mean free path appropriate to the gas at the upper boundary of the slab, and throughout the slab the density is regarded as constant. When the upper boundary is raised, the slab thickens somewhat, but even so the lower boundary is also raised with the result that however high the level considered, the formula never gives the complete loss from all parts of the atmosphere.

To obtain the total loss of molecules from an atmosphere by the present method, we have now to integrate the expression (7·09) over all values of $r$ from $r_c$ (the critical level) to infinity, but in doing so we must remember that $C_r$ may be one of two different functions of $r$. As pointed out in the preceding section, there is a stratum in which $\mathscr{C}_r > c_r$, while above it $c_r > \mathscr{C}_r$. So that in one case $C_r$ is to be associated with $\mathscr{C}_r$ and in the other with $c_r$. At the boundary of the two regions $\mathscr{C}_r$ is equal to $c_r$. This height we shall denote by $r_d$. Its value is determined by the equations

$$\frac{\pi^{\frac{1}{2}}\mathscr{C}_r^2 hm}{\psi(\mathscr{C}_r\sqrt{hm})} = \frac{e^{q_0\frac{r_0}{r_d}} - 1}{e^{q_0\frac{r_0}{r_c}} - 1}, \quad \text{.................................(7·10)}$$

and

$$\mathscr{C}_r^2 = c_r^2 = 2g\frac{a^2}{r_d} \quad \text{.......................................(7·11)}$$

* Bromwich, *Theory of Infinite Series*, p. 326.

† Jeans, *loc. cit.*, p. 343.

The equation for the height of $r_c$ has already been given in equation (4·08), but for convenience it can be repeated here*.

$$e^{q_0\frac{r_0}{r_c}} - 1 = \frac{q_0 e^{q_0}}{\pi\nu_0\sigma^2 r_0}. \qquad \text{.......(7·12)}$$

In the region $r_c$ to $r_d$, we are to substitute for $C_r$ in (7·09), before integrating, its value given by

$$\frac{\sqrt{\pi}\, hm\, C_r^2}{\psi(C_r\sqrt{hm})} = \frac{e^{q_0\frac{r_0}{r}} - 1}{e^{q_0\frac{r_0}{r_d}} - 1}, \qquad \text{.......(7·13)}$$

while above $r_d$, we get the simple relation

$$C_r^2 = c_r^2 = 2ga^2/r. \qquad \text{.......(7·14)}$$

In formulae (7·11) and (7·14), $a$ is the radius of the planet, but this may be equated to the value of $r_0$ appropriate to the base of the isothermal part of the atmosphere. (In the case of the earth, $a = 6370$ kilometres, while $r_0 = 6390$ kilometres.)

It is convenient first to calculate the loss from above $r_d$. Making the necessary substitutions for $\nu$, $c_r$ and $\mathscr{C}_r$, this is given by

$$L_d = \frac{4\pi^{\frac{3}{2}}\sigma^2\nu_0^2 r_0^4 e^{-2q_0}}{(hm)^{\frac{1}{2}}} \int_{r_d}^{\infty} \frac{e^{q_0\frac{r_0}{r}}}{r^2} \left\{\tfrac{3}{2} - 2\pi^{\frac{1}{2}}\lambda\left(e^{q_0\frac{r_0}{r}} - 1\right) + \frac{q_0 r_0}{r} - \lambda q_0 \frac{r_0}{r}\left(e^{q_0\frac{r_0}{r}} - 1\right) + \frac{\lambda r}{q_0 r_0}\left(e^{q_0\frac{r_0}{r}} - 1\right)\right\} dr$$

$$= \frac{4\pi^{\frac{3}{2}}\sigma^2\nu_0^2 r_0^4 e^{-2q_0}}{(hm)^{\frac{1}{2}} q_0 r_0} \left[e^{q_0\frac{r_0}{r_d}}\left\{\tfrac{1}{2} + \frac{q_0 r_0}{r_d} + \lambda\left(1 + q_0\frac{r_0}{r_d}\right)\right\} - \lambda e^{2q_0\frac{r_0}{r_d}}\left(\tfrac{3}{4} + \frac{q_0 r_0}{2r_d}\right) - \tfrac{1}{2} - \lambda\left(\tfrac{1}{4} - \frac{q_0 r_0}{r_d}\right)\right]. \qquad \text{.......(7·15)}$$

None of the integrals calls for comment except

$$\int_{r_d}^{\infty} \frac{e^{2q_0\frac{r_0}{r}} - e^{q_0\frac{r_0}{r}}}{r}\, dr.$$

When an obvious change of variable is made, this is seen to consist of the difference of two logarithmic integrals, each of which diverges but whose difference remains finite. For large values of $r_d$ (and it is always large), we find that the integral is equal approximately to $q_0\frac{r_0}{r_d}$.

For brevity, we have put

$$\lambda = \frac{\pi\nu_0\sigma^2 r_0}{q_0 e^{q_0}} = \frac{1}{e^{q_0\frac{r_0}{r_c}} - 1}. \qquad \text{.......(7·16)}$$

It is large only when $r_c$ is large compared with $q_0 r_0$, that is when there is appreciable density at heights above the surface of the planet equal to several times its radius. Clearly there is no

* In the earlier part of the work we used an approximate formula for the distribution of atmospheric density, given by Milne. This is a sufficiently close representation of the actual distribution both in the very high and in the very low parts of the atmosphere when $q_0$ is large. In the general problem, now being investigated, which we hope to be applicable to all planets, $q_0$ may be small (see footnote, page 548). In such cases, Milne's formula, while sufficiently accurate for most purposes, does not give the critical level with sufficient closeness. And so, in the present application, we introduce a new approximation which removes the objection just mentioned and at the same time gives a close representation of the actual distribution in the parts of the atmosphere with which we are mainly concerned. The formula in question is

$$\nu = \nu_0\left(\frac{r_0}{r}\right)^2 e^{-q_0\left(1-\frac{r_0}{r_c}\right)},$$

the $(q_0 - 2)$ of Milne's formula being replaced by $q_0$.

need to discuss such cases. In the planetary atmospheres as we know them, $r_c$ is of the same order as $r_c$ and so $\lambda$ will be a very small quantity. (In the case of the earth $r_c$ is largest for hydrogen, and even in that case we have seen above that it is equal only to $1{\cdot}238\, r_0$; the corresponding value of $\lambda$ is $10^{-24}$.) We have retained all the terms containing $\lambda$ so that the formula for $L_d$ might be applied to cases, if there be any, where $\lambda$ is no longer small. For our present purpose, we may now write

$$L_d = \frac{4\pi^{\frac{3}{2}}\sigma^2\nu_0^2 r_0^4 e^{-2q_0}}{(hm)^{\frac{1}{2}} q_0 r_0}\left[e^{q_0\frac{r_0}{r_d}}\left(\tfrac{1}{2}+\frac{q_0 r_0}{r_d}\right) - \lambda e^{2q_0\frac{r_0}{r_d}}\left(\tfrac{3}{4}+\frac{q_0 r_0}{2r_d}\right)\right]. \quad \text{............(7·17)}$$

As we have seen, this expression does not give the complete loss. We have still to calculate the escape from the region $r_c$ to $r_d$, viz.

$$\frac{4\pi^{\frac{3}{2}}\sigma^2\nu_0^2 r_0^4 e^{-2q_0}}{(hm)^{\frac{1}{2}}}\int_{r_c}^{r_d}\frac{e^{2q_0\frac{r_0}{r}}}{r^2}\, e^{-hm\mathscr{C}_r^2}$$
$$\times\left\{\left(\tfrac{3}{2} - \frac{2\mathscr{C}_r^2 hm\sqrt{\pi}}{\psi(\mathscr{C}_r\sqrt{hm})}\right) + \sqrt{\pi}\,\mathscr{C}_r^2 hm\left(1 - \frac{\mathscr{C}_r^2 hm}{\psi(\mathscr{C}_r\sqrt{hm})}\right) + \frac{\sqrt{\pi}}{4\psi(\mathscr{C}_r\sqrt{hm})}\right\} dr,$$

where $\mathscr{C}_r$ is a function of $r$ determined by equation (7·13). This relation is so complicated as to preclude any hope of effecting the integration by analysis. Its value can only be obtained accurately in particular cases by quadrature. It is very easy, however, to get an idea of its order of magnitude. For since in this region $\mathscr{C}_r > c_r$, the value of the above integral is less than it would be if $\mathscr{C}_r$ were replaced by $c_r$. It follows that the actual loss is less than $L_c$—the expression obtained by changing $r_c$ to $r_d$ wherever it occurs in (7·17)—while it is greater than $L_d$. Now

$$\frac{L_d}{L_c} = \frac{r_c}{r_d}\,\frac{e^{q_0\frac{r_0}{r_d}}}{e^{q_0\frac{r_0}{r_c}}}\,\frac{2-\lambda e^{q_0\frac{r_0}{r_d}}}{2-\lambda e^{q_0\frac{r_0}{r_c}}}$$

very nearly. But by definition, equation (7·16),

$$\lambda e^{q_0\frac{r_0}{r_c}} = 1+\lambda$$
$$= 1$$

very approximately, so that

$$\frac{L_d}{L_c} = \frac{r_c}{r_d}\, e^{q_0\left(\frac{r_0}{r_d}-\frac{r_0}{r_c}\right)}\left\{2 - e^{q_0\left(\frac{r_0}{r_d}-\frac{r_0}{r_c}\right)}\right\},$$

which clearly depends primarily on the ratio of $e^{q_0\frac{r_0}{r_d}}$ to $e^{q_0\frac{r_0}{r_c}}$. This we can easily find, for we have seen above, equations (7·10) and (7·11), that

$$\frac{\pi^{\frac{1}{2}} q_0 r_0/r_d}{\psi(\sqrt{q_0 r_0/r_d})} = \frac{e^{q_0\frac{r_0}{r_d}}-1}{e^{q_0\frac{r_0}{r_c}}-1}$$
$$= e^{q_0 r_0\left(\frac{1}{r_d}-\frac{1}{r_c}\right)} \quad \text{..............................(7·18)}$$

approximately. The value of the left-hand side of this equation is easily obtained in particular cases from Table II, on page 542. It is small only when $q_0 r_0/r_d$ is small. A rough calculation shows that if hydrogen existed on the moon with a molecular density at the surface of $10^7$ per unit volume, $q_0 r_0/r_d$ is slightly less than 3·93 and the expressions (7·18) are equal approximately to ·5.

In such a case the region $r_c$ to $r_d$ supplies at most a number of molecules equal to those which escape from above $r_d$. We shall accordingly regard $L_c$ as giving with sufficient accuracy the loss in all other cases. So finally, we have, changing $r_d$ to $r_c$ in equation (7·17) and substituting for λ,

$$L = \frac{4\pi^{\frac{3}{2}}\sigma^2\nu_0^2 r_0^4 e^{-2q_0} e^{q_0\frac{r_0}{r_c}}}{(hm)^{\frac{1}{2}} q_0 r_0}\left(\frac{q_0 r_0}{2r_c} - \tfrac{1}{4}\right)$$

$$= \frac{2\pi^{\frac{3}{2}}\sigma^2\nu_0^2 r_0^4 e^{-2q_0} e^{q_0\frac{r_0}{r_c}}}{(hm)^{\frac{1}{2}} r_c}. \quad\text{.................................(7·19)}$$

We may suppose that the statical molecular distribution, which is disturbed by the escape of these molecules, is restored by processes of diffusion so that the loss occasioned by the evaporation of molecules from the region above $r_c$ is made good by the diffusion of an equal number of molecules across the critical layer.

If for any reason the level $r_c$ could be regarded as fixed, equation (7·19) shows that the rate of loss would be proportional to the square of the molecular concentration at the base. In general, however, the height of this critical level is itself a function of $\nu_0$ and will gradually diminish as a result of the escape of molecules. The relation between it and $\nu_0$ has been given above, equation (7·12), and when substituted in the expression for $L$ we find a remarkable simplification. Thus

$$L = \frac{2\pi^{\frac{1}{2}} q_0 r_0^3 e^{-q_0}\nu_0}{(hm)^{\frac{1}{2}} r_c} = \frac{4(\pi hm)^{\frac{1}{2}} g r_0^4 e^{-2hmgr_0}\nu_0}{r_c}, \quad\text{..................(7·20)}$$

which not only is independent of the size of the molecules but depends only on the first power of $\nu_0$. In calculating the time for an atmosphere to stream away, to which we proceed in the next section, we shall require both formulae, (7·19) and (7·20), for $L$.

§ 8. *The Time to lose a Simple Atmosphere.* In an ideal atmosphere, such as we are now considering, which consists of one gas only, the critical level $r_c$ will continue to descend until it reaches the surface of the planet. As we have seen the rate of loss of the atmosphere during this stage is given by equation (7·20) and is proportional to the *first power* of the molecular density at the base, but the critical height having become fixed, the formula (7·19) must be used and this involves the *square* of the density.

In the first stage we have, if $N$ be the total number of molecules present,

$$\frac{dN}{dt} = -\frac{4(\pi hm)^{\frac{1}{2}} g r_0^4 e^{-2hmgr_0}}{r_c}\nu_0. \quad\text{.........................(8·01)}$$

Now $N$ is given by $$N = \int_{r_0}^{\infty} 4\pi r^2 \nu\, dr = \nu_0 4\pi r_0^2 e^{-q_0}\int_{r_0}^{\infty} e^{q_0\frac{r_0}{r}}\, dr.$$

This is a divergent integral, but the difficulty is only formal and can be avoided by considering the atmosphere to have a ceiling of large radius $R$. We then get (using the first term of an asymptotic expression)

$$N = 4\pi\nu_0 r_0^2 e^{-q_0}\left[-e^{q_0\frac{r_0}{R}}\frac{r^2}{q_0 r_0}\right]_{r_0}^{R}$$

$$= \frac{4\pi\nu_0 r_0^3}{q_0}. \quad\text{..............................................(8·02)}$$

Following Jeans*, we can get $N$ by a very simple method, thus: suppose the atmosphere to

* Jeans, *loc. cit.*, p. 345.

be equivalent to a layer of thickness $H$ of uniform molecular density $\nu_0$. Then at the base, the pressure of the gas is $\nu_0 mgH$. But it is also $\nu_0/2h$. Hence $H=\frac{1}{2hmg}=\frac{r_0}{q_0}$ and so $N=4\pi r_0^2 H$ as above.

The equation (8·01) can accordingly be written

$$\frac{d\nu_0}{dt}=-\frac{q_0^2 e^{-q_c}}{2(\pi hm)^{\frac{1}{2}} r_c}\nu_0=-k_1\nu_0, \qquad (8\cdot03)$$

which gives the law of diminution of $\nu_0$ in the first stage of escape. Although $k_1$ contains $r_c$, it may with sufficient accuracy be regarded as constant. It is only when $r_c$ occurs in an exponential term that it is important.

Suppose that at time $t=0$, the value of $\nu_0$ is $(\nu_0)_0$ and the corresponding value of the critical level is $r_c$. The corresponding value of $\nu_0$ when the critical level reaches the surface of the planet is easily seen from equation (7·12) to be

$$\frac{(\nu_0)_0}{(\nu_0)'}=\frac{e^{q_0}-1}{e^{q_0\frac{r_0}{r_c}}-1}=e^{q_0\left(1-\frac{r_0}{r_c}\right)} \qquad (8\cdot04)$$

approximately. The value of $\nu_0$ at any time is clearly

$$\nu_0=(\nu_0)_0 e^{-k_1 t}, \qquad (8\cdot05)$$

and so the time for the change from $(\nu_0)_0$ to $(\nu_0)'$ is

$$T_1=\frac{1}{k_1}\log_e\frac{(\nu_0)_0}{(\nu_0)'}=\frac{q_0}{k_1}\left(1-\frac{r_0}{r_c}\right)$$

$$=\frac{2(\pi hm)^{\frac{1}{2}} r_c e^{q_0}}{q_0}\left(1-\frac{r_0}{r_c}\right)$$

$$=\frac{\sqrt{2\pi}\, r_c e^{\frac{3gr_0}{C^2}} C}{\sqrt{3gr_0}}\left(1-\frac{r_0}{r_c}\right), \qquad (8\cdot06)$$

if $C$ is the average velocity of thermal agitation of the gas.

In the second stage of the escape we have

$$\frac{dN}{dt}=-\frac{2\pi^{\frac{3}{2}}\sigma^2 r_0^3 e^{-q_0}}{(hm)^{\frac{1}{2}}}\nu_0^2, \qquad (8\cdot07)$$

which gives

$$\frac{d\nu_0}{dt}=-\frac{\pi^{\frac{1}{2}}\sigma^2 e^{-q_0}}{2(hm)^{\frac{1}{2}}}\nu_0^2$$

$$=-k_2\nu_0^2, \qquad (8\cdot08)$$

say.

The escape now proceeds much more slowly, for we have

$$\frac{1}{\nu_0}-\frac{1}{(\nu_0)'}=k_2 t, \qquad (8\cdot09)$$

and so the time to change from $(\nu_0)'$ to $(\nu_0)''$ is

$$T_2=\frac{1}{k_2}\left\{\frac{1}{(\nu_0)''}-\frac{1}{(\nu_0)'}\right\}. \qquad (8\cdot10)$$

If $(\nu_0)''$ is one $n$th part of $(\nu_0)'$, we can write

$$T_2=\frac{1}{k_2}\frac{n}{\nu_0'} \qquad (8\cdot11)$$

if $n$ is large. The time now depends directly on the ratio $n$ of the densities, whereas in the first stage it depended only on the logarithm of $n$.

For example, suppose that the atmosphere of Mars consisted only of helium and that at a certain time in its history the molecular concentration at the surface of the planet was $10^{15}$. Then supposing the temperature to be $-100^\circ$ C., the time required for the critical layer to reach the surface of the planet is $6{\cdot}98 \times 10^9$ years, the corresponding time for a temperature of $0^\circ$ C. being $3{\cdot}88 \times 10^4$ years. At the end of this period the basic molecular concentrations are $6{\cdot}6 \times 10^7$ and $4{\cdot}17 \times 10^7$ respectively. Afterwards the dissipation takes place much more slowly. It is easily seen that at the lower temperature it takes a further $1{\cdot}4 \times 10^{10} n$ years to reduce the density at the base to one $n$th of $6{\cdot}6 \times 10^7$. Various numerical values are given in the table at the end of the next section, where the effect of other atmospheric constituents is considered.

§ 9. *The Loss of a Mixed Atmosphere.* Actually an atmosphere consists of one or more constituent gases and, according to the law formulated by Dalton, each is distributed as though it alone were present. We have already given a mathematical expression for the law of distribution. It is clear that the density falls off least rapidly in the case of the lightest gas, and so at very great heights the atmosphere consists of this constituent alone. Its loss by escape proceeds as in the case of a simple atmosphere just discussed if its critical level is at such a height that the density of the other gases present is there negligible. In the course of time, however, this critical level, in its gradual descent, will enter the region of the next heaviest and will then sink more slowly. At the same time, the character of the escape will change for the light gas will escape not only owing to collisions among its own molecules but also owing to collisions with the molecules of the second gas. The net effect of these two considerations requires investigation.

The work of § 2 on the chance of a collision of one molecule of a gas with any other can easily be extended to a mixture of gases. Denoting as before the chance in time $\delta t$ by $\Theta(c)\,\delta t$, we find in the case of two gases

$$\Theta(c) = \Theta_{11}(c) + \Theta_{12}(c), \quad \ldots\ldots(9{\cdot}01)$$

where

$$\Theta_{12}(c) = \frac{\sqrt{\pi}\,\nu_2 \sigma_{12}^2\, \psi(c\sqrt{hm_2})}{chm_2}. \quad \ldots\ldots(9{\cdot}02)$$

The symbol $\sigma_{12}$ denotes the sum of the radii of the two different kinds of molecules. Arguing as before, the length of a free path is seen to be given by

$$\frac{\sqrt{\pi}\,\sigma_{11}^2\,\psi(c\sqrt{hm_1})}{c^2hm_1}\int \nu_1\,ds + \frac{\sqrt{\pi}\,\sigma_{12}^2\,\psi(c\sqrt{hm_2})}{c^2hm_2}\int \nu_2\,ds = 1. \quad \ldots\ldots(9{\cdot}03)$$

Using the law of distribution

$$\nu_1 = (\nu_1)_0 \left(\frac{r_0}{r}\right)^2 e^{-q_1\left(1-\frac{r_0}{r}\right)}, \quad \ldots\ldots(9{\cdot}04)$$

with a corresponding expression for $\nu_2$, we find for a molecule moving from a point $r$ in a direction making an angle $\theta$ with the vertical,

$$\frac{\sqrt{\pi}\,\sigma_{11}^2\,\psi(c\sqrt{hm_1})\,(\nu_1)_0\,e^{-q_1 r_0}}{c^2hm_1\,q_1}\left[e^{q_1\frac{r_0}{r}} - e^{q_1\frac{r}{R}}\right] + \frac{\sqrt{\pi}\,\sigma_{12}^2\,\psi(c\sqrt{hm_2})\,(\nu_2)_0\,e^{-q_2 r_0}}{c^2hm_2\,q_2}\left[e^{q_2\frac{r_0}{r}} - e^{q_2\frac{r}{R}}\right] = \cos\theta.$$

The critical level for the first gas is obtained by putting $c = \infty$, $R = \infty$, and $\theta = 0$. The equation becomes

$$\pi\sigma_{11}^2(\nu_1)_0\frac{e^{-q_1 r_0}}{q_1}\left[e^{q_1\frac{r_0}{r_c}} - 1\right] + \pi\sigma_{12}^2(\nu_2)_0\frac{e^{-q_2 r_0}}{q_2}\left[e^{q_2\frac{r_0}{r_c}} - 1\right] = 1. \quad \ldots\ldots(9{\cdot}05)$$

The unit terms in the brackets are in nearly all cases small in comparison with the other terms and can be neglected. The relation can then also be written in the form

$$\frac{\sigma_{11}^2 (\nu_1)_c}{m_1} + \frac{\sigma_{12}^2 (\nu_2)_c}{m_2} = \frac{2hgr_0^2}{r_c^2 \pi}, \quad \ldots\ldots(9\cdot06)$$

where $(\nu_1)_c$ and $(\nu_2)_c$ are the respective molecular densities at the critical level.

The escape of molecules from a spherical shell of radius $r$ is then given by

$$dL = 4\pi r^2 \nu_1 f_1 (u, v, w)\, 2\pi c^2 \sin\theta\, d\theta\, dc\, dr\, \{\Theta_{11}(c) + \Theta_{12}(c)\},$$

and, using the methods of § 7, we then find

$$dL = dL_{11} + dL_{12},$$

where $dL_{11}$ is the same as $dL$ given in equation (7·03) and

$$dL_{12} = \frac{8\pi\sigma_{12}^2 (hm_1)^{\frac{3}{2}} \nu_1 \nu_2 r^2 dr}{hm_2} \int_{C_r}^{\infty} e^{-hm_1c^2} c\, \psi(c\sqrt{hm_2})(1-\cos\theta)\, dc.$$

It can be shown that when this expression is integrated from the critical height $r_c$ to infinity,

$$L_{12} = \frac{2\pi^{\frac{3}{2}} \sigma_{12}^2 (hm_1)^{\frac{1}{2}} (\nu_1)_0 (\nu_2)_0 r_0^4 e^{-q_1} e^{-q_2}}{hm_2\, q_2 r_0} \left[ e^{q_2 \frac{r_0}{r_c}} \left( q_2 \frac{r_0}{r_c} - 1 \right) + 1 + \frac{m_1 + m_2}{m_1} \left( e^{q_2 \frac{r_0}{r_c}} - 1 \right) \right]$$

$$= \frac{2\pi^{\frac{3}{2}} \sigma_{12}^2 (hm_1)^{\frac{1}{2}} (\nu_1)_0 (\nu_2)_0 r_0^4 e^{-q_1} e^{-q_2} e^{q_2 \frac{r_0}{r_c}}}{hm_2 r_c} \quad \ldots\ldots(9\cdot07)$$

approximately. Hence

$$L = \frac{2\pi^{\frac{3}{2}} r_0^4 (\nu_1)_0 (hm_1)^{\frac{1}{2}} e^{-q_1}}{hr_c} \left[ \frac{\sigma_{11}^2 (\nu_1)_c}{m_1} + \frac{\sigma_{12}^2 (\nu_2)_c}{m_2} \right] \frac{r_c^2}{r_0^2} \quad \ldots\ldots(9\cdot08)$$

We have already found a relation (equation (9·06)) for the function in the brackets and so we find a remarkable simplification,

$$L = \frac{4 (\pi hm)^{\frac{1}{2}} r_0^4 (\nu_1)_0\, g e^{-q_1}}{r_c} \quad \ldots\ldots(9\cdot09)$$

This is precisely the same expression obtained for the escape of a simple atmosphere when the critical level is sinking. In this case the critical level will never become fixed until the heavier gases of the atmosphere have themselves streamed away to such an extent that the total density is so small that the critical level for all constituents is at the surface of the planet. The escape of the lightest element may be regarded as given by equation (9·09) throughout its entire escape, and is proportional at all times to the first power of the basic molecular concentration.

As in the preceding section, we thus find

$$\frac{d\nu_0}{dt} = -k_1 \nu_0$$

as the law of escape, with the same meaning of $k_1$ as before. The time to reduce the density to one $n$th of its initial value is therefore given by

$$T = \frac{\log_e n}{k_1}.$$

The loss of helium from Mars, assuming the presence of heavier constituents such as water vapour, has been calculated according to this law and the results are compared in the following table with the loss supposing it existed there alone.

TABLE III.

*The Loss of Helium from Mars. (Time in years.)*

| $\nu_0$ | $-100^\circ$ C. | | $0^\circ$ C. | |
|---|---|---|---|---|
| | Simple Atmosphere | Mixed Atmosphere | Simple Atmosphere | Mixed Atmosphere |
| $10^{15}$ | 0 | 0 | 0 | 0 |
| $10^{12}$ | $2{\cdot}97 \times 10^{9}$ | $2{\cdot}97 \times 10^{9}$ | $1{\cdot}58 \times 10^{4}$ | $1{\cdot}58 \times 10^{4}$ |
| $10^{9}$ | $5{\cdot}94 \times 10^{9}$ | $5{\cdot}94 \times 10^{9}$ | $3{\cdot}16 \times 10^{4}$ | $3{\cdot}16 \times 10^{4}$ |
| $6{\cdot}6 \times 10^{7}$ | $6{\cdot}98 \times 10^{9}$ | $6{\cdot}98 \times 10^{9}$ | | |
| $4{\cdot}17 \times 10^{7}$ | | | $3{\cdot}88 \times 10^{4}$ | $3{\cdot}88 \times 10^{4}$ |
| $10^{6}$ | $9{\cdot}77 \times 10^{11}$ | $8{\cdot}91 \times 10^{9}$ | $2{\cdot}08 \times 10^{6}$ | $4{\cdot}74 \times 10^{4}$ |
| $10^{5}$ | $9{\cdot}77 \times 10^{12}$ | $1{\cdot}19 \times 10^{10}$ | $2{\cdot}08 \times 10^{7}$ | $6{\cdot}32 \times 10^{4}$ |
| $10^{4}$ | $9{\cdot}77 \times 10^{13}$ | $1{\cdot}49 \times 10^{10}$ | $2{\cdot}08 \times 10^{8}$ | $7{\cdot}90 \times 10^{4}$ |
| $10^{3}$ | $9{\cdot}77 \times 10^{14}$ | $1{\cdot}79 \times 10^{10}$ | $2{\cdot}08 \times 10^{9}$ | $9{\cdot}48 \times 10^{4}$ |

In the same way, the loss of the earth's atmosphere can be considered. If we suppose hydrogen to exist with a basic concentration* of $1{\cdot}89 \times 10^{13}$, then at a temperature of $-54^\circ$ C., the number which escape per second is $1{\cdot}5 \times 10^{8}$. Its basic concentration will be reduced to $1{\cdot}89 \times 10^{8}$ at the end of $2 \times 10^{24}$ years. The corresponding values for the temperature of $27^\circ$ C., recently given by Lindemann and Dobson†, are $1{\cdot}67 \times 10^{16}$ molecules per second and $2{\cdot}68 \times 10^{16}$ years.

It is evident that according to the present method the time of escape of an atmosphere is longer in all cases than that given by Jeans. For example, the times given by him for the complete loss of helium from Mars at the temperatures $-100^\circ$ C. and $0^\circ$ C. are respectively $10^{9}$ and $10^{3}$ years. All atmospheres retained according to Jeans' calculations will a fortiori be retained according to the present method and so his main conclusions concerning the constitution of the atmospheres of the planets of the solar system are unaffected. We may likewise suppose that the results obtained by Milne‡ in the case of non-isothermal atmospheres by the more approximate method would be borne out by the methods of this paper.

In conclusion, the author wishes to acknowledge his indebtedness to Mr R. H. Fowler, with whom he has had the privilege of discussing this paper in its various stages and from whom he has received many helpful suggestions.

* Chapman and Milne, *loc. cit.*, see also p. 541 above.

† Lindemann and Dobson, *Proc. Roy. Soc.*, Jan. 1923.

‡ Milne, *loc. cit.*

# GENERAL INDEX

CAMBRIDGE: PRINTED BY W. LEWIS, AT THE UNIVERSITY PRESS